



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 102834

TO: Nita M Minnifield
Location: CM1/8A07&8E12
Art Unit: 1645
_____, 2003

Case Serial Number: 10/030231

From: P. Sheppard
Location: CM1-1E03
Phone: (703) 308-4499

sheppard@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

102834

From: Chan, Christina
Sent: Tuesday, September 02, 2003 11:44 AM
T : Minnifield, Nita; STIC-Biotech/ChemLib
Subject: RE: rush sequence search

RECEIVED

SEP -2 2003

STIC-BIOTECH/ChemLib
(STIC)

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: Minnifield, Nita
Sent: Tuesday, September 02, 2003 11:40 AM
To: Chan, Christina
Subject: rush sequence search

Christina, please approve, 2 month amdt.

STIC
10/030231

Please do a commercial sequence search on SEQ ID NO: 5 and 6 (amino acid sequences) of this application.

Please provide a paper copy of the results.

Thanks,

Nita M. Minnifield
Art Unit 1645
Office CM1-8A07
Mailbox CM1-8E12
703-305-3394

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:44:21 ; Search time 13.5 Seconds

(without alignments)
139,338 Million cell updates/sec

Title: US-10-030-231-5

Perfect score: 199
Sequence: 1 HXHXTSYXCKFCGTAXCTYXCRXLHGKXCXCHCSR 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	177	88.9	96 1 MYNA_MYTGA	P8103 mytilus gal
2	175	87.9	96 1 MYNB_MYTGA	P82102 mytilus gal
3	53.5	26.9	595 1 2440_HUMAN	Q81y18 homo sapien
4	52.5	26.4	186 1 AG13_WHEAT	P10969 triticum ae
5	52.5	26.4	212 1 AG1_HORVU	P13112 hordeum vul
6	52	26.1	372 1 VLP1_CAEEL	Q03571 caenorhabdi
7	51.5	25.9	72 1 IBB_VICAN	P01065 vicia angus
8	49.5	24.9	1271 1 TRPS_XENLA	Q90286 xenopus lae
9	49	24.6	194 1 KRUB_HUMAN	Q75690 homo sapien
10	49	24.6	352 1 TX37_CAEEL	Q942C9 caenorhabdi
11	49	24.6	379 1 CYR6_MOUSE	P18406 mus musculu
12	49	24.6	379 1 CYR6_MOUSE	Q96872 rattus norv
13	48.5	24.4	1339 1 ADO_BOVIN	P48034 bos taurus
14	47.5	23.9	458 1 2275_HUMAN	Q9nsd4 homo sapien
15	47.5	23.9	517 1 2215_HUMAN	Q9u158 homo sapien
16	47.5	23.9	1801 1 LMB2_RAT	P15800 rattus norv
17	47	23.6	261 1 UL31_HSVGA	Q01041 herpesvirus
18	46.5	23.4	213 1 AG12_WHEAT	P02876 triticum ae
19	46.5	23.4	534 1 2397_HUMAN	Q8hf99 homo sapien
20	46.5	23.4	806 1 TRB7_MOUSE	P26011 mus musculu
21	46.5	23.4	1281 1 TRPS_HUMAN	Q9uhf7 homo sapien
22	46.5	23.4	1281 1 TRPS_HUMAN	Q925h1 mus musculu
23	46.5	23.4	1895 1 WR19_ARATH	Q94267 arabidopsis
24	46.5	23.4	1895 1 REIN_HUMAN	P78509 homo sapien
25	46.5	23.4	3461 1 REIN_MOUSE	Q06841 mus musculu
26	46.5	23.4	3462 1 REIN_MOUSE	P58751 rattus norv
27	46.5	23.4	3579 1 STRN_DROME	Q9v5n8 drosophila
28	46.5	23.4	3712 1 LMA_DROME	Q00174 drosophila
29	46	23.1	343 1 NOV_XENLA	P51609 xenopus lae
30	46	23.1	381 1 CYR6_HUMAN	Q00622 homo sapien
31	46	23.1	488 1 U711_HUMAN	Q9wuf9 homo sapien
32	46	23.1	532 1 ZN20_HUMAN	P17024 homo sapien
33	46	23.1	606 1 Z666_XENLA	P18733 xenopus lae

34	45.5	22.9	577	1	ITB6_CANPO	P18563 cavia porce
35	45.5	22.9	677	1	SP87_DICDI	P54643 dictyosteli
36	45.5	22.9	788	1	ITB6_HUMAN	P18564 homo sapien
37	45.5	22.9	792	1	2328_HUMAN	Q8wxb4 homo sapien
38	45.5	22.9	798	1	ITB5_MOUSE	O70309 mus musculu
39	45.5	22.9	1955	1	AGRI_CHICK	P31696 gallus galli
40	45	22.6	84	1	EC1_ARATH	P93746 arabidopsis
41	45	22.6	169	1	KRU4_HUMAN	P26371 homo sapien
42	45	22.6	218	1	YIDX_ECOLI	P31461 escherichia
43	45	22.6	860	1	IDLH_HUMAN	P01130 homo sapien
44	45	22.6	1172	1	TSP2_MOUSE	Q03350 mus musculu
45	45	22.6	1428	1	ATRN_MOUSE	Q9wu60 mus musculu

ALIGNMENTS

RESULT 1

MYNA_MYTGA STANDARD: PRT: 96 AA.
ID MYNA_MYTGA
AC P82103;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mytilin A precursor.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytilidae;
OC Mytilidae; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-56.
RC TISSUE=Hemocyte, and Plasma;
RX MEDLINE=99421718; PubMed=10491159;
RA Mitra G., Hubert F., Noel T., Roch P.;
RT "Mytilin, a novel cysteine-rich antimicrobial peptide isolated from
hemocytes and plasma of the mussel Mytilus galloprovincialis.";
RL Eur. J. Biochem. 265:71-78(1999).
CC -!- FUNCTION: BACTERIOLYTIC ACTIVITY AGAINST GRAM-POSITIVE BACTERIA
M. LUTEUS, B. MEGATERIUM AND A. VIRIDANS.
CC -!- TISSUE SPECIFICITY: HEMOCYTES.
CC -!- PTM: FOUR DISULFIDE BONDS ARE PRESENT.
CC -!- MASS SPECTROMETRY: MW=4437.45; METHOD=Electrospray.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: AF162334; AADA7638.1; -
KW SIGNAL; Antibiotic.
FT SIGNAL 1 20
FT PEPTIDE 21 60 MYTICIN A.
FT PROPEP 61 96 REMOVED IN MATURE FORM.
SQ SEQUENCE 96 AA; 10556 MW; 0C7B1CC35E5C2B0 CRC64;

Query Match 88.9%; Score 177; DB 1; Length 96;
Best Local Similarity 70.0%; Pred. No. 7.9e-18;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

0Y 1 HXHXTSYXCKFCGTAXCTYXCRXLHGKXCXCHCSR 40
| | | | | | | | | | | | | | | | | | | | | |
Db 21 HSHACTSYWCKFCGTACTGHTLCLVLRGKMKACAVHCSR 60

RESULT 2

MYNB_MYTGA STANDARD: PRT: 96 AA.
ID MYNB_MYTGA
AC P82102;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

DE Mytilin B precursor.

OS Mytilus galloprovincialis (Mediterranean mussel).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;

OC Mytiloidae; Mytilidae; Mytilus.

NCBI_Taxid=29158;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-27.

RC TISSUE=Hemocyte;

RA MEDLINE=99421718; PubMed=10491159;

RA Malta G., Hubert F., Noel T., Roch P.;

RT "Mytilin, a novel cysteine-rich antimicrobial peptide isolated from hemocytes and plasma of the mussel Mytilus galloprovincialis.";

RL Eur. J. Biochem. 255:71-78(1999).

CC -1- FUNCTION: BACTERIOLYTIC ACTIVITY AGAINST GRAM-POSITIVE BACTERIA

CC M. LUTENS, B. MEGATERIUM AND A. VIRIDANS AND GRAM-NEGATIVE BACTERIA

CC E. COLI D31. POSSESSES ANTI-FUNGAL ACTIVITY AGAINST F. OXYSPORUM.

CC -1- TISSUE SPECIFICITY: HEMOCYTES.

CC -1- PPM: FOUR DISULFIDE BONDS ARE PRESENT.

CC -1- MASS SPECTROMETRY: MW=4563.45; METHOD=Electrospray.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC DR EMBL: AF162335; AAD47639.1; -

CC KW Antibiotic; Fungicide; Signal.

CC FT SIGNAL 1 20 MYTICIN B.

CC FT CHAIN 21 60 REMOVED IN MATURE FORM.

CC FT PROPEP 61 96

CC FT SEQUENCE 96 AA; 10653 MW; FB084F33D10E9232 CRC64;

CC SQ

Query Match 87.9%; Score 175; DB 1; Length 96;

Best Local Similarity 70.0%; Pred. No. 1.5e-17;

Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1 HXHXCTGYCKKFGCTACTYXXCXRLXGKXCKXHCGR 40

DB 21 HPHVCTSYCKKFGCTACTYXXCXRLXGKXCKXHCGR 60

RESULT 3

2440. HUMAN STANDARD: PRT; 595 AA.

AC Q8Y1B8; Q8Y1B8; 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Zinc finger protein 440.

GN ZNF440.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Tongue;

RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,

RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,

RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,

RA Kawachi H., Saito K., Nishikawa T., Kimura K., Yamashita H.,

RA Matsuo K., Nakamura K., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,

RA Murakami K., Kanehori K., Takahashi-Fujii A., Oshino A., Sugiyama M.,

RA Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K.,

RA Isogai T.;

RT "NEO human cDNA sequencing project.";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udell T.B., Toshyuki S., Carninci P., Mulhany S.J.,

RA Raba S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,

RA Schnerch A., Schein J.E., Jones S.J.W., Maira W.A.,

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- FUNCTION: May be involved in transcriptional regulation.

CC -1- SUBCELLULAR LOCATION: Nuclear (potential).

CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.

CC -1- SIMILARITY: Contains 1 KRAB domain.

CC -1- SIMILARITY: Contains 12 C2H2-type zinc fingers.

CC -----

CC DR EMBL: BC035760; AAH35760.1; -

CC DR EMBL: AK095252; BAC04510.1; -

CC DR Genew: HGNC:20874; ZNF440.

CC DR InterPro: IPR001909; KRAB.

CC DR InterPro: IPR007087; Znf_C2H2.

CC DR Pfam: PF01352; KRAB; 1.

CC DR Pfam: PF00096; zf_C2H2; 11.

CC DR SMART: SM00349; KRAB; 1.

CC DR SMART: SM00355; Znf_C2H2; 11.

CC DR PROSITE: PS00805; KRAB; 1.

CC DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 10.

CC DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 12.

CC DR Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;

CC KW Nuclear protein; Repeat.

CC FT DOMAIN 4 86 KRAB

CC FT ZN_FING 144 166 C2H2-TYPE 1 (DEGENERATE).

CC FT ZN_FING 172 194 C2H2-TYPE 2.

CC FT ZN_FING 200 222 C2H2-TYPE 3.

CC FT ZN_FING 228 250 C2H2-TYPE 4.

CC FT ZN_FING 256 278 C2H2-TYPE 5 (DEGENERATE).

CC FT ZN_FING 284 306 C2H2-TYPE 6.

CC FT ZN_FING 312 334 C2H2-TYPE 7.

CC FT ZN_FING 340 362 C2H2-TYPE 8.

CC FT ZN_FING 368 390 C2H2-TYPE 9.

CC FT ZN_FING 396 418 C2H2-TYPE 10.

CC FT ZN_FING 424 448 C2H2-TYPE 11.

CC FT ZN_FING 456 480 C2H2-TYPE 12.

CC FT ZN_FING 488 510 L -> M (IN REF. 1).

CC FT CONFLICT 40 569 S -> N (IN REF. 1).

CC FT CONFLICT 569 595 7DCDDFB698E40661 CRC64;

CC SQ SEQUENCE 595 AA; 69105 MW; 7DCDDFB698E40661 CRC64;

Query Match 26.9%; Score 53.5; DB 1; Length 595;

Best Local Similarity 34.1%; Pred. No. 2.1; 11; Indels 13; Gaps 4;

Matches 14; Conservative 3; Mismatches 11; Indels 13; Gaps 4;

OY 8 YXCKKFGCTACTYXXC-----RXLXHG-KXCKXHCGR 40

28-FEB-2003 (Rel. 41, Last annotation update)

Root-specific lectin precursor.

OS Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;

OC Triticeae; Hordeum.

OC NCBI_Taxid=4513;

RM [1]

RP SEQUENCE FROM N.A.

RA Lerner D. R., Ralkehl N. V.;

RT "Cloning and characterization of root-specific barley lectin.";

RU Plant Physiol. 91:124-129(1989).

CC -1- FUNCTION: CARBOHYDRATE BINDING.

CC -1- SIMILARITY: Contains 4 chitin-binding domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>

CC or send an email to license@isb-sdb.ch).

CC -----

CC EMBL: M29280; AAA32969.1; .

CC DR PIR: T05936; T05936.

CC DR HSSP: P10969; 1MGP.

CC DR Interpro: IPR001002; Chitin_binding_1.

CC DR Pfam: PF00187; Chitin_bind.1; 4.

CC DR PRINTS: PR00451; CHITINBINDNG.

CC DR SMART: SM00270; CnCB1; 4.

CC DR PROSITE: PS00026; CHITIN_BINDING: 4.

CC KW Lectin; Repeat; Chitin-binding; Glycoprotein; Signal;

CC PYRrolidone carboxylic acid.

CC KW SIGNAL 1 26

CC FT CHAIN 27 212

CC FT DOMAIN 27 69

CC FT DOMAIN 70 112

CC FT DOMAIN 113 155

CC FT DOMAIN 156 197

CC FT DOMAIN 27 27

CC MOD_RES 27

CC FT DISULFID 29 44

CC FT DISULFID 38 50

CC FT DISULFID 43 57

CC FT DISULFID 61 66

CC FT DISULFID 72 87

CC FT DISULFID 81 93

CC FT DISULFID 104 100

CC FT DISULFID 104 109

CC FT DISULFID 115 130

CC FT DISULFID 124 136

CC FT DISULFID 129 143

CC FT DISULFID 147 152

CC FT DISULFID 158 173

CC FT DISULFID 167 179

CC FT DISULFID 172 186

CC FT DISULFID 190 195

CC FT CARBOHYD 206

CC SQ SEQUENCE 212 AA; 21209 MW; 8D948245D6B25A5 CRC64;

Query Match 26.4%; Score 52.5; DB 1; Length 212;

Best Local Similarity 30.2%; Pred. No. 1.2;

Matches 13; Conservative 1; Mismatches 20; Indels 9; Gaps 2.2;

QY 5 CTST-----XCXKFCGTACX-TYXCRXLHGKXCKXCHC 38

DB 44 CSYGYCGMGDYCGKCGQNGACYSRCGTQAGSKTCPPNNHC 86

RESULT 6

Y1FL CAEEL

ID Y1FL CAEEL

AC 003571;

STANDARD;

PRT; 372 AA.

```

DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical RNA-binding protein C40H1.1 in chromosome III.
GN      C40H1.1.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitae; Rhabditoidea;
OC      Caenoditidae; Peioderinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-Bristol NZ;
RC      MEDLINE=94150718; PubMed=7906398;
RA      Wilson R., Alnsough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA      Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
RA      Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA      Fulton L., Jones A., Kerhaw J., Kirsten J., Laister N.,
RA      Latreille P., Lighting J., Lloyd C., Mortimore B., O'Callaghan M.,
RA      Parsons J., Percy C., Rifken L., Roopra A., Saunders E., Staden R.,
RA      Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Vaughan K.,
RA      Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA      Waterston R., Watson A., Weinstock L., Wilkinson-Sprat J.,
RA      Wohlfman P.;
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans.";
CC      -1- SIMILARITY: Contains 2 RNA recognition motif (RNM) domains.
CC      -1- SIMILARITY: NO DIPOPHYLIA OVARIAN PROTEIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: Z19154; CAAT9552.1; -.
CC      PIR: S28296; S28296.
DR      WormPep: C40H1.1; CE00109.
DR      InterPro: IPR000504; RNA_rec_mot.
DR      Pfam: PF00076; rrm; 2.
DR      SMART: SM00360; RRM; 2.
DR      PROSITE: PS50102; RRM; 2.
DR      PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
DR      ProSite: PS00030; RRM_RNP_1; FALSE_NEG.
DR      Hypothetical protein: RNA-binding. Repeat.
FT      DOMAIN          58      161      RNA-BINDING (RRM) 1.
FT              178      249      RNA-BINDING (RRM) 2.
FT      DOMAIN          178      249      RNA-BINDING (RRM) 2.
SQ      SEQUENCE       372 AA: 42393 MW; 8016734389D97E18 CRC64;
QY      Query Match           26.1%; Score 52; DB 1; Length 372;
DB      Best local Similarity 43.5%; Pred. No. 2.3;
Matches 10; Conservative 1; Mismatches 8; Indels 4; Gaps 1;
        13 FCGTACTCTYXCR---XLHXGK 31
        ||||| | | | | : ||
IBR_VICAN
ID      IBR_VICAN         STANDARD;             PRT;             72 AA.
AC      P01065;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      21-JUL-1998 (Rel. 36, Last annotation update)
DE      Bowman-Birk type proteinase inhibitor (VBI).
OS      Vicia angustifolia (Common vetch).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX      NCBI_TaxID=3909;

```



```

RN [1]
RP SEQUENCE.
RC STRAIN=cv. Segetalis:
RA Shinozuka Y., Kuromitzu K., Araki T., Ohata J., Abe O.;
RT "The complete amino acid sequence of Vicia angustifolia proteinase
inhibitor."
RL Nat. Cult. 10:69-73(1983).
CC -1- FUNCTION: THIS INHIBITOR HAS TWO DOMAINS, EACH WITH SEPARATE
CC ANTI-PROTEASE ACTIVITY. 1 MOLE OF INHIBITOR INHIBITS EITHER 1 MOLE
CC OF TRYPSIN OR 2 MOLES OF CHYMOTRYPSIN, SPOICHOIMERICALLY.
CC -1- SIMILARITY: BELONGS TO THE BOWMAN-BIRK SERINE PROTEASE INHIBITORS
CC FAMILY.
CC PIR: A01304; TITVOA.
CC DR HSSP: P56679; IPBI.
CC DR InterPro: IPR000877; Bowman-Birk Leg.
CC DR Pfam: PF00228; Bowman-Birk_Leg; 2.
CC DR ProDom: PD002168; Bowman-Birk_Leg; 1.
CC DR SMART: SM00269; BowB; 1.
CC DR PROSITE: PS00281; BOWMAN_BIRK; 1.
CC KW Serine protease inhibitor.
CC FT ACT_SITE 16 17 INTERACTION WITH TRYPSIN.
CC FT ACT_SITE 42 43 INTERACTION WITH CHYMOTRYPSIN.
CC FT DISULFID 8 61 BY SIMILARITY.
CC FT DISULFID 9 24 BY SIMILARITY.
CC FT DISULFID 12 57 BY SIMILARITY.
CC FT DISULFID 14 22 BY SIMILARITY.
CC FT DISULFID 31 38 BY SIMILARITY.
CC FT DISULFID 35 50 BY SIMILARITY.
CC FT DISULFID 40 48 BY SIMILARITY.
CC FT DISULFID 72 AA; 8038 MW; 21AADB275DE33D83 CRC64;
SQ
Query Match 25.9%; Score 51.5; DB 1; Length 72;
Best Local Similarity 36.7%; Pred. No. 0.69; Mismatches 12; Indels 5; Gaps 2;
Matches 11; Conservative 2;
OY 14 CGTACTYX--XCRXLHXGXC--XCXHC 38
DB 9 CDTCLCTRSQPTCRVDCVGERCHSACNHC 38

```

```

CC -1- SIMILARITY: Contains 9 C2H2-type zinc fingers.
CC -1- SIMILARITY: Contains 1 GATA-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF346838; AAK39510.1; -
CC DR EMBL: AF346837; AAK39509.1; -
CC DR InterPro: IPR007087; znf.C2H2.
CC DR InterPro: IPR000679; znf.GATA.
CC DR Pfam: PF00320; GATA; 1.
CC DR PRINTS: PR00096; zt-C2H2; 3.
CC DR SMART: SM00355; znf.C2H2; 9.
CC DR SMART: SM00401; znf.GATA; 1.
CC DR PROSITE: PS00344; GATA_ZN_FINGER_1; 1.
CC DR PROSITE: PS0114; GATA_ZN_FINGER_2; 1.
CC DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
CC DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; FALSE_NEG.
CC KW Transcription regulation; Repressor; Zinc-finger; DNA-binding;
CC Nuclear protein; Metal-binding; Repeat; Alternative splicing.
CC FT ZN_FING 217 242 C2H2-TYPE 1 (ATYPICAL).
CC FT ZN_FING 328 353 C2H2-TYPE 2 (ATYPICAL).
CC FT ZN_FING 426 451 C2H2-TYPE 3 (ATYPICAL).
CC FT ZN_FING 513 543 C2H2-TYPE 4 (ATYPICAL).
CC FT ZN_FING 604 627 C2H2-TYPE 5.
CC FT ZN_FING 656 679 C2H2-TYPE 6.
CC FT ZN_FING 682 705 C2H2-TYPE 7.
CC FT ZN_FING 886 910 GATA-TYPE.
CC FT ZN_FING 1205 1227 C2H2-TYPE 8.
CC FT ZN_FING 1205 1257 C2H2-TYPE 9.
CC FT VARSPLIC 364 616 Missing (in isoform mTRPS1).
CC FT CONFLICT 848 848 A -> S (IN REF. 1; AAK39509).
CC FT CONFLICT 852 852 K -> R (IN REF. 1; AAK39509).
CC SQ SEQUENCE 1271 AA; 141436 MW; BFA6841A939E1DCA CRC64;
Query Match 24.9%; Score 49.5; DB 1; Length 1271;
Best Local Similarity 31.1%; Pred. No. 14;
Matches 14; Conservative 1; Mismatches 17; Indels 13; Gaps 3;
OY 3 HXCTSYXCKFCGTACTYXCRXLHXG-----KXCKXCHCS 39
DB 547 HKCTIKHC-PFCPRGLCT-----PEKHLGEITYFPACKKSCNSHCA 586

```

```

RESULT 8
TRPS_XENLA STANDARD; PRT; 1271 AA.
ID TRPS_XENLA Q90ZS6;
AC Q90ZS6; Q90ZS7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger transcription factor Trps1.
GN TRPS1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8335;
RN
RP SEQUENCE FROM N.A. (ISOFORMS TRPS1 AND MTRPS1), AND CHARACTERIZATION.
RX MEDLINE-21181829; PubMed-1128535;
RA Malik T.H., Shoichet S.A., Latham P., Kroll T.G., Peters L.L.,
RA Shivasani R.A.;
RT "Transcriptional repression and developmental functions of the
RT atypical vertebrate GATA protein Trps1."
RL EMO J. 20:1715-1725(2001).
CC -1- FUNCTION: Transcriptional repressor. May act to restrict
CC expression of GATA-regulated genes at selected sites and stages in
CC vertebrate development.
CC -1- SUBUNIT: Binds specifically to GATA sequences.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing: Named isoforms-2;
CC Name-xTrps1; Synonyms-zygotic;
CC IsoId-Q90ZS6-1; Sequence-Displayed;
CC Name-mTRPS1; Synonyms-Maternal;
CC IsoId-Q90ZS6-2; Sequence-VSP_001607;

```

```

RESULT 9
KRUB_HUMAN STANDARD; PRT; 194 AA.
ID KRUB_HUMAN Q75690;
AC Q75690;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE keratin, ultra high-sulfur matrix protein B (UHS keratin B) (UHS
DE kerat).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE-99148005; PubMed-10023043;
RA Perez C., Aurio J., Gerst C., Bernard B.A., Egly J.-M.;
RT Genomic organization and promoter characterization of two human UHS
RT keratin genes."
RL Gene 227:137-148(1999).
CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID

```

DR PIR; T31544; T31544.
DR HSSP; P24781; 1XBR.
DR WormPep; Y47D3A.12; CE24328.
DR InterPro; IPR001699; TF_T-box
DR PIR; T31544; T31544.

```
CC -----
DR EMBL; M32490; AAA37512.1; -
DR EMBL; X56790; CAA40109.1; -
DR PIR; A35669; A35669.
```

DR MCD; MGI:88613; Cyt6L.
 DR GO; GO:0001569; P:pattern of blood vessels; IMP.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot_1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR Pfam: PF00093; VWC; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; VWF_C; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 DR Growth factor binding; Signal.
 KM SIGNAL 1 24
 FT CHAIN 1 24
 FT DOMAIN 25 379
 FT DOMAIN 98 164
 FT DOMAIN 226 271
 FT DOMAIN 284 358
 FT DISULFID 284 321
 FT DISULFID 301 335
 FT DISULFID 312 351
 FT DISULFID 315 353
 FT DISULFID 320 357
 SQ SEQUENCE 379 AA; 41709 MW; FA6B5014B56AAEE9 CRC64;
 Query Match 24.6%; Score 49; DB 1; Length 379;
 Best Local Similarity 28.0%; Pred. No. 5.9;
 Matches 14; Conservative 4; Mismatches 22; Indels 10; Gaps 3;
 OY 1 HXHKC---TSYX-CXKFCGTAXCXYYX-----CRXLHGKXCXCHCSR 40
 DB 223 HGQKCIQVTTSMQSCSGGTGISTRTVNDNSECHLVKTRICEVRPCGQ 272
 RESULT 12
 CYR6_RAT STANDARD; PRT; 379 AA.
 AC Q9E572;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cyt6L protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (Insulin-like growth factor-binding protein 10).
 GN Cyt6L OR IGFBP10.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=20435857; PubMed=10852911;
 RA Albrecht C., von der Kammer H., Mayhaus M., Klaudny J., Schweizer M.,
 Nitsch R.M.;
 RT "Muscarinic acetylcholine receptors induce the expression of the
 RT immediate early growth regulatory gene Cyt6L.";
 RT J. Biol. Chem. 275:28929-28936(2000).
 CC -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS: PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Contains 1 IGFBP domain.
 CC -!- SIMILARITY: Contains 1 VWF domain.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.

CC -!- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF218568; AAG14964.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot_1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR Pfam: PF00093; VWC; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; VWF_C; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 DR Growth factor binding; Signal.
 KM SIGNAL 1 24
 FT CHAIN 1 24
 FT DOMAIN 25 379
 FT DOMAIN 98 164
 FT DOMAIN 226 271
 FT DOMAIN 284 358
 FT DISULFID 284 321
 FT DISULFID 301 335
 FT DISULFID 312 351
 FT DISULFID 315 353
 FT DISULFID 320 357
 SQ SEQUENCE 379 AA; 41687 MW; 62BF0BBA4C5AFDE9 CRC64;
 Query Match 24.6%; Score 49; DB 1; Length 379;
 Best Local Similarity 28.0%; Pred. No. 5.9;
 Matches 14; Conservative 4; Mismatches 22; Indels 10; Gaps 3;
 OY 1 HXHKC---TSYX-CXKFCGTAXCXYYX-----CRXLHGKXCXCHCSR 40
 DB 223 HGQKCIQVTTSMQSCSGGTGISTRTVNDNSECHLVKTRICEVRPCGQ 272
 RESULT 13
 ADO_BOVIN STANDARD; PRT; 1339 AA.
 AC P48034;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Aldehyde oxidase (EC 1.2.3.1).
 GN AOX1 OR AO.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 43-48; 187-218 AND 538-573.
 RC TISSUE=Liver;
 RX MEDLINE=96125080; PubMed=8537361;
 RA Calz M.L., Ravitolo C., Ghisaldi E., de Gioia L., Salmona M.,
 Ca zaniga G., Kurosaki M., Terao M., Garattini E.;

RT "Purification, cDNA cloning, and tissue distribution of bovine liver
 RT aldehyde oxidase." J. Biol. Chem. 270:31037-31045(1995).
 CC J. Biol. Chem. 270:31037-31045(1995).
 CC -1 CATALYTIC ACTIVITY: An aldehyde + H(2)O + O(2) -> a carboxylic acid
 CC + H(2)O(2).
 CC -1 COFACTOR: FAD, MOLYBDOPTERIN, AND TWO 2FE-2S CLUSTERS.
 CC -1 SUBUNIT: Homodimer.
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1 TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG AND
 CC SPLEEN.
 CC -1 PTM: THE N-TERMINUS IS BLOCKED.
 CC -1 SIMILARITY: BELONGS TO THE XANTHINE DEHYDROGENASE FAMILY.
 CC -1 SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 CC EMBL: X87251; CAA60701.1; -
 CC HSSP: P80457; 1FC4.
 DR InterPro: IPR002888; 2Fe-2S_bind.
 DR InterPro: IPR006058; 2Fe2S_ferredoxin.
 DR InterPro: IPR000674; Aldxan_dh.C.
 DR InterPro: IPR005107; CO_deh_flav.C.
 DR InterPro: IPR002346; dehydrog_molyb.
 DR InterPro: IPR000572; Euk_Mb_oxred.
 DR InterPro: IPR001041; Ferredoxin.
 DR Pfam: PF02738; Ald_Xan_dh.C2; 1.
 DR Pfam: PF03450; CO_deh_flav.C; 1.
 DR Pfam: PF00941; FAD_binding_5; 1.
 DR Pfam: PF00111; fer2; 1.
 DR Pfam: PF01799; fer2_2; 1.
 DR ProDom: PD186071; 2Fe-2S_bind; 1.
 DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
 DR PROSITE: PS00559; MOLYBDOPTERIN_EUK; 1.
 DR Oxidoreductase: NAD; Molybdenum; Flavoprotein; FAD; Metal-binding;
 KM Iron-sulfur; Iron; 2Fe-2S.
 FT METAL 44 44 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 49 49 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 52 52 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 SQ SEQUENCE 1339 AA; 147610 MW; 3CA7FFPD2806655 CRC64;
 Query Match 24.4%; Score 48.5; DB 1; Length 1339;
 Best local Similarity 32.4%; Pred. No. 20; Indels 9; Gaps 2;
 Matches 11; Conservative 2; Mismatches 12;
 Db 5 CTSY-----XCXKFCGTAXCTYXCRXLHXGXC 33
 151 CTGYRPIINACKTFCFTSGC-----CQSKENGVCC 180
 RESULT 14
 Z275_HUMAN STANDARD; PRT; 458 AA.
 ID Z275_HUMAN
 AC Q9NSD4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 275.
 GN ZNF275.
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCB1_TaxID=9606;
 RN NCB1 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20314869; PubMed=10854409; Gloeckner G., Botcherby M.,
 RA Mallon A.M., Platzer M., Bates R.,

RA Nordstiek G., Strivens M.A., Kioschis P., Dangel A., Cunningham D.,
 RA Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K.,
 RA Kerry G., Greystrong J.S., Clark D., Goerdes M., Blechschmidt K.,
 RA Rump A., Hinzmann B., Mundy C.R., Miller W., Poustka A., Herman G.E.,
 RA Rhodes W., Denny P., Rosenthal A., Brown S.D.M.;
 RT "Comparative genome sequence analysis of the Bpa/Stt region in mouse
 RT and man.";
 RL Genome Res. 10:758-775(2000).
 CC -1 FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1 SUBCELLULAR LOCATION: Nuclear (Probable)
 CC -1 SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -1 SIMILARITY: Contains 1 KRAB domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 CC EMBL: U82670; AAF4786.1; -
 CC HSSP: P08047; 1SP2.
 DR Genew: HGNC:13069; ZNF275.
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam: PF01352; KRAB; 1.
 DR Pfam: PF00096; zf-C2H2; 1.
 DR ProDom: PD000003; Znf_C2H2; 4.
 DR SMART: SM00349; KRAB; 1.
 DR SMART: SM00355; Znf_C2H2; 11.
 DR PROSITE: PS50805; KRAB; 1.
 DR PROSITE: PS50028; ZINC_FINGER_C2H2_1; 11.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 11.
 DR Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KM Nuclear protein; Repeat; Repressor.
 KM DOMAIN 6 71 KRAB.
 FT ZN_FING 130 152 C2H2-TYPE.
 FT ZN_FING 158 180 C2H2-TYPE.
 FT ZN_FING 210 232 C2H2-TYPE.
 FT ZN_FING 238 260 C2H2-TYPE.
 FT ZN_FING 266 288 C2H2-TYPE.
 FT ZN_FING 294 316 C2H2-TYPE.
 FT ZN_FING 322 344 C2H2-TYPE.
 FT ZN_FING 350 372 C2H2-TYPE.
 FT ZN_FING 378 400 C2H2-TYPE.
 FT ZN_FING 406 428 C2H2-TYPE.
 FT ZN_FING 434 457 C2H2-TYPE.
 SQ SEQUENCE 458 AA; 52024 MW; A2ADED98AF791779 CRC64;
 Query Match 23.9%; Score 47.5; DB 1; Length 458;
 Best local Similarity 35.1%; Pred. No. 11; Indels 5; Gaps 3;
 Matches 13; Conservative 2; Mismatches 17;
 Db 8 YXKRCGTAXCTYXCC--RXLHXG-KXGXCHCSR 40
 406 YEC DK-CGKAFRRSSGLSHRRHISGARCESSQGR 441
 RESULT 15
 Z215_HUMAN STANDARD; PRT; 517 AA.
 ID Z215_HUMAN
 AC Q9UL58;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 215 (BMSCR2 associated zinc-finger protein 2) (BAZ
 DE 2).
 GN ZNF215 OR BAZ2.
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20311529; PubMed=10762538;
RA Alders M., Ryan A., Hodges M., Bilek J., Feinberg A.P., Privitera O.,
RA Westerveld A., Little P.F.R., Mannens M.;
RT "Disruption of a novel imprinted zinc-finger gene, ZNF215, in
RT Beckwith-Wiedemann syndrome.";
RL Am. J. Hum. Genet. 66:1473-1484(2000).
CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 1 SCAN box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF056618; AAF00005.1; -.
DR HSSP; P08047; ISP1.
DR Genev; HGNC:13007; ZNF215.
DR MIM; 605016; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0008270; F:zinc ion binding activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003309; Treg_SCAN.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF02023; SCAN; 1.
DR Pfam; PF00096; zf_C2H2; 4.
DR ProDom; PD000003; Znf_C2H2; 4.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00431; LER; 1.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS50804; SCAN_BOX; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein.
KW DOMAIN 48 126 SCAN_BOX.
FT DOMAIN 164 237 KRAB.
FT ZN_FING 379 401 C2H2-TYPE.
FT ZN_FING 407 429 C2H2-TYPE.
FT ZN_FING 462 484 C2H2-TYPE.
FT ZN_FING 490 512 C2H2-TYPE.
SQ SEQUENCE 517 AA; 60048 MM; 33569C960E6D3F19 CRC64;

Query Match 23.9%; Score 47.5; DB 1; Length 517;
Best Local Similarity 34.2%; Pred. No. 12;
Matches 13; Conservative 3; Mismatches 17; Indels 5; Gaps 3;

QY 7 SYCXKFCGTAXCTYYXC---RXLHXG-KXCXCHCSR 40
Db 378 SYECYQ-CGKAFCRSSSLIRHQIITHGKPYKCSGCR 414

Search completed: September 3, 2003, 10:59:31
Job time : 15.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:44:21; Search time 13.5 seconds

(without alignments)
139,338 Million cell updates/sec

Title: US-10-030-231-6

Perfect score: 252
Sequence: 1 HSHACTSYWCGKFCGTASCTHYLCRVLHPGRKACVHCSR 40

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	252	100.0	96	1 MYNA_MYTGA	P82103 mytilus gal
2	180	71.4	96	1 MYNB_MYTGA	P82102 mytilus gal
3	65	23.8	461	1 PRCG_MOUSE	P33587 mus musculu
4	61.5	24.4	1202	1 JAG2_RAT	P97607 rattus norv
5	61.5	24.4	1247	1 JAG2_MOUSE	O9945 mus musculu
6	61	24.2	459	1 PRCG_PIG	O94122 sus scrofa
7	60.5	24.0	71	1 NKX1_NAJHH	P25674 najia haje h
8	59.5	23.6	1238	1 JAG2_HUMAN	O99219 homo sapien
9	59	23.4	71	1 NKX1_NAJUNI	P01390 najia nivea
10	59	23.4	111	1 ZG32_XENLA	P18719 xenopus lae
11	59	23.4	1210	1 AT19_MOUSE	P59509 mus musculu
12	58.5	23.2	212	1 AG1_HORVU	P15312 hordeum vul
13	58.5	23.2	965	1 YNC3_YEAST	P53971 saccharomyc
14	58.5	23.2	1242	1 JAG1_BRARE	O90577 brachydantio
15	58.5	23.2	5147	1 FAT_DROME	P33450 drosophila
16	58	23.0	1895	1 WR19_ARATH	O95267 arabidopsis
17	58	23.0	3075	1 LMA1_HUMAN	P25391 homo sapien
18	57	22.6	180	1 TR22_MOUSE	O96962 mus musculu
19	56.5	22.4	186	1 AG13_WHEAT	P10869 triticum ae
20	56.5	22.4	410	1 T801_PSESH	P24607 pseudomonas
21	56	22.2	72	1 NKX1_NAJHC	P01389 najia haje a
22	56	22.2	182	1 KRCU_SHEEP	O03672 ovis aries
23	56	22.2	372	1 YLF1_CAEEL	O03571 caenorhabdi
24	56	22.2	458	1 PRCG_RABIT	O28651 cryotolagus
25	56	22.2	461	1 PRCG_HUMAN	P04070 homo sapien
26	56	22.2	461	1 PRCG_RAT	P31394 rattus norv
27	56	22.2	517	1 Z215_HUMAN	O9u158 homo sapien
28	56	22.2	838	1 RN19_HUMAN	O9u558 homo sapien
29	56	22.2	840	1 RN19_MOUSE	P50636 mus musculu
30	55.5	22.0	71	1 NKX2_NAJME	P01388 najia melano
31	55.5	22.0	956	1 AD19_HUMAN	O9u013 homo sapien
32	55	21.8	148	1 LVC_CERAE	P30200 ceratophyc
33	55	21.8	148	1 LVC_COLAN	P79698 colobus ang

34	55	21.8	148	1 LVC_MACMU	P30201 macaca mula
35	55	21.8	148	1 LVC_MIOTA	P79806 mlopithecus
36	55	21.8	148	1 LVC_PAPAN	P00696 papio anubi
37	55	21.8	379	1 CYR6_MOUSE	P18406 mus musculu
38	55	21.8	379	1 CYR6_RAT	O98872 rattus norv
39	55	21.8	551	1 YK77_YEAST	P36113 saccharomyc
40	55	21.8	815	1 AD15_MOUSE	O88839 mus musculu
41	54	21.4	148	1 LVC_HYLLA	P79180 hylobates 1
42	54	21.4	272	1 YWV3_CAEEL	O11077 caenorhabdi
43	54	21.4	488	1 U713_HUMAN	O9u5f9 homo sapien
44	54	21.4	2813	1 VWF_HUMAN	P04275 homo sapien
45	54	21.4	3066	1 POL6_BCMVN	O65399 b genome po

ALIGNMENTS

RESULT 1	ID	MYNA_MYTGA	STANDARD:	PRT:	96 AA.
AC	P82103:	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Mytilin A precursor.				
OS	Mytilus galloprovincialis (Mediterranean mussel).				
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;				
OC	Mytiloidea; Mytilidae; Mytilus.				
OX	NCBI_Taxid=29158;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 21-56.				
RC	TISSUE=hemocyte, and Plasma;				
RA	MEDLINE=99421718; PubMed=10491159;				
RA	Mitra G., Hubert F., Noel T., Roch P.;				
RT	"Mytilin, a novel cysteine-rich antimicrobial peptide isolated from hemocytes and plasma of the mussel Mytilus galloprovincialis.";				
RT	Eur. J. Biochem. 265:71-78(1999).				
RL	- FUNCTION: BACTERIOLYTIC ACTIVITY AGAINST GRAM-POSITIVE BACTERIA				
CC	M. LUTEUS, B. MEGATERIUM AND A. VIRIDANS.				
CC	- TISSUE SPECIFICITY: HEMOCYTES.				
CC	- PTM: FOUR DISULFIDE BONDS ARE PRESENT.				
CC	- MASS SPECTROMETRY: MW=4437.45; METHOD=Electrospray.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
DR	EMBL: AF162334; AAD47638.1; -				
FT	SIGNAL; Antibiotic. 20				
FT	PEPTIDE 21 60				
FT	PROPEP 61 96				
SO	SEQUENCE 96 AA: 10556 MW: 007B1C3535C2B0 CRC64;				
QY	Query Match	100.0%;	Score 252;	DB 1;	Length 96;
	Best Local Similarity	100.0%;	Pred. No. 6.2e-23;		
	Matches 40;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
DB	1 HSHACTSYWCGKFCGTASCTHYLCRVLHPGRKACVHCSR 40				
	21 HSHACTSYWCGKFCGTASCTHYLCRVLHPGRKACVHCSR 60				
RESULT 2					
MYNB_MYTGA	STANDARD:	PRT:	96 AA.		
AC	P82102:				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				

DE 28-FEB-2003 (Rel. 41, last annotation update)
 DT Myticin B precursor.
 OS Mytilus galloprovincialis (Mediterranean mussel).
 CC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
 CC Mytilidae; Mytilidae; Mytilus.
 CC NCBI_TaxID=29158;
 RN (1)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-27.
 RC TISSUE=Hemocyte;
 RX MEDLINE=99421718; PubMed=10491159;
 RA Mita G., Hubert F., Noel T., Roch P.;
 "Myticin, a novel cysteine-rich antimicrobial peptide isolated from
 hemocytes and plasma of the mussel Mytilus galloprovincialis.",
 Eur. J. Biochem. 265:71-78(1999).
 CC -1- FUNCTION: BACTERIOLYTIC ACTIVITY AGAINST GRAM-POSITIVE BACTERIA
 M. LUTERUS, B. MEGATERIUM AND A. VIRIDANS AND GRAM-NEGATIVE BACTERIA
 E. COLI D31. POSSESSES ANTIFUNGAL ACTIVITY AGAINST F. OXYSPORUM.
 CC -1- TISSUE SPECIFICITY: HEMOCYTES.
 CC -1- TISSUE SPECIFICITY: HEMOCYTES.
 CC -1- PTM: FOUR DISULFIDE BONDS ARE PRESENT.
 CC -1- MASS SPECTROMETRY: MW=4563.45; METHOD=Electrospray.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF162335; AADA7639.1; -
 KW Antibiotic; Fungicide; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 60 MYTICIN B.
 FT PROPEP 61 96 REMOVED IN MATURE FORM.
 SQ SEQUENCE 96 AA; 10653 MW; FB084F3D10E9232 CRC64;
 Query Match 71.4%; Score 180; DB 1; Length 96;
 Best local Similarity 70.0%; Pred. No. 1.4e-14;
 Matches 28; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

RESULT 3

PRIC_MOUSE STANDARD; PRT: 461 AA.
 ID PRTIC_MOUSE
 AC P33587; 035498;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
 DE (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
 factor XIV).
 DE PROC.
 OS Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92316897; PubMed=1618739;
 RA Tada N., Sato M., Tsujimura A., Iwase R., Hashimoto-Gotoh T.;
 "Isolation and characterization of a mouse protein C cDNA.",
 J. Biochem. 111:491-495(1992).
 RL (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvJ;
 RX MEDLINE=98155576; PubMed=9493582;
 RA Jalbert L.R., Rosen E.D., Lissens A., Carmeliet P., Collen D.,
 Castellino F.J.;

RT "Nucleotide structure and characterization of the murine gene encoding
 RT anticoagulant protein C.";
 RL Thromb. Haemost. 79:310-316(1998).
 RN (3)
 RP SEQUENCE OF 274-434 FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=94318474; PubMed=8043441;
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
 "A comparative study of partial primary structures of the catalytic
 RT region of mammalian protein C.";
 RL Br. J. Haematol. 86:590-600(1994).
 CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
 CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
 CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
 CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
 CC and VIIIA.
 CC -1- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
 CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
 CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
 CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
 CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
 CC STRONGLY PROMOTED BY THROMBOMODULIN.
 CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
 CC -1- PTM: THE VITAMIN K-DEPENDENT ENZYMAIC CARBOXYLATION OF SOME
 CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
 CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
 CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
 CC SITE IS NECESSARY FOR THE RECOGNITION OF THE
 CC THROMBIN-THROMBOMODULIN COMPLEX.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D10445; BAA01235.1; -
 DR EMBL: AF034569; AAC3795.1; -
 DR EMBL: D43755; BAA07812.1; -
 DR PIR: JX0210; JX0210.
 DR HSSP: P04070; IPCU.
 DR MEROPS: S01.218; -
 DR MGD: MGI:97771; PROC.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR006209; EGF_Like.
 DR InterPro: IPR002383; GLA_blood.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR InterPro: IPR000294; Vitk_dep_GLA.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00594; gla; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00001; GLABLOOD.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00069; GLA; 1.
 DR SMART: SM00020; TRYP_SPE; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
 DR PROSITE: PS00240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR Blood coagulation; Glycoprotein; Serine protease;
 KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
 KW EGF-like domain; Repeat; Endothelial cell; Hydrolyase; Signal.


```

FT DOMAIN 452 488 EGF-LIKE 8.
FT DOMAIN 490 527 EGF-LIKE 9.
FT DOMAIN 529 589 EGF-LIKE 10 (ATYPICAL).
FT DOMAIN 591 627 EGF-LIKE 11, CALCITON-BINDING (POTENTIAL).
FT DOMAIN 629 665 EGF-LIKE 12, CALCITON-BINDING (POTENTIAL).
FT DOMAIN 667 703 EGF-LIKE 13.
FT DOMAIN 706 742 EGF-LIKE 14.
FT DOMAIN 744 780 EGF-LIKE 15, CALCITON-BINDING (POTENTIAL).
FT DOMAIN 782 818 EGF-LIKE 16, CALCITON-BINDING (POTENTIAL).
FT DISULFID 199 210 BY SIMILARITY.
FT DISULFID 203 216 BY SIMILARITY.
FT DISULFID 218 227 BY SIMILARITY.
FT DISULFID 230 241 BY SIMILARITY.
FT DISULFID 236 247 BY SIMILARITY.
FT DISULFID 249 258 BY SIMILARITY.
FT DISULFID 265 277 BY SIMILARITY.
FT DISULFID 271 287 BY SIMILARITY.
FT DISULFID 289 298 BY SIMILARITY.
FT DISULFID 305 316 BY SIMILARITY.
FT DISULFID 327 336 BY SIMILARITY.
FT DISULFID 343 354 BY SIMILARITY.
FT DISULFID 348 363 BY SIMILARITY.
FT DISULFID 365 374 BY SIMILARITY.
FT DISULFID 381 392 BY SIMILARITY.
FT DISULFID 386 401 BY SIMILARITY.
FT DISULFID 403 412 BY SIMILARITY.
FT DISULFID 419 429 BY SIMILARITY.
FT DISULFID 423 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 456 467 BY SIMILARITY.
FT DISULFID 461 476 BY SIMILARITY.
FT DISULFID 478 487 BY SIMILARITY.
FT DISULFID 495 506 BY SIMILARITY.
FT DISULFID 500 515 BY SIMILARITY.
FT DISULFID 517 526 BY SIMILARITY.
FT DISULFID 544 567 POTENTIAL.
FT DISULFID 561 577 POTENTIAL.
FT DISULFID 579 588 BY SIMILARITY.
FT DISULFID 595 606 BY SIMILARITY.
FT DISULFID 600 615 BY SIMILARITY.
FT DISULFID 617 626 BY SIMILARITY.
FT DISULFID 633 644 BY SIMILARITY.
FT DISULFID 638 653 BY SIMILARITY.
FT DISULFID 655 664 BY SIMILARITY.
FT DISULFID 671 682 BY SIMILARITY.
FT DISULFID 676 691 BY SIMILARITY.
FT DISULFID 693 702 BY SIMILARITY.
FT DISULFID 710 721 BY SIMILARITY.
FT DISULFID 715 730 BY SIMILARITY.
FT DISULFID 732 741 BY SIMILARITY.
FT DISULFID 746 759 BY SIMILARITY.
FT DISULFID 753 768 BY SIMILARITY.
FT DISULFID 770 779 BY SIMILARITY.
FT DISULFID 786 797 BY SIMILARITY.
FT DISULFID 791 806 BY SIMILARITY.
FT DISULFID 808 817 BY SIMILARITY.
SQ SEQUENCE 1202 AA; 129703 MW; 08CB44E5271FF8BF CRC64;

Query Match 24.48; Score 61.5; DB 1; Length 1202;
Best Local Similarity 31.08; Pred. No. 7.3;
Matches 18; Conservative 1; Mismatches 16; Indels 23; Gaps 3;

```

```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN Jagged 2 precursor (Jagged2).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Swiss Webster / NIH;
RS Tsai S.;
RN Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RC [2]
RX SEQUENCE OF 302-819 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98051918; PubMed=9341252;
RA Lan Y., Jiang R., Shawber C., Weinmaster G., Gridley T.;
RT "The Jagged2 gene maps to chromosome 12 and is a candidate for the 191
RL Mamm. Genome 8:875-876(1997).
RN [3]
RP SEQUENCE OF 325-759 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98145947; PubMed=9486542;
RA Valasech C., Ghezzi C., Ballabio A., Rugari E.I.;
RT "Jagged2, a putative Notch ligand expressed in the apical ectodermal
RL ridge and in sites of epithelial-mesenchymal interactions.";
RN Mech. Dev. 69:203-207(1997).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=97459705; PubMed=9315665;
RA Luo B., Aster J.C., Hasseljian R.P., Kuo F., Sklar J.;
RT "Isolation and functional analysis of a cDNA for human Jagged2, a gene
RL encoding a ligand for the Notch receptor.";
RN Mol. Cell. Biol. 17:6057-6067(1997).
RN [5]
RP FUNCTION: PUTATIVE NOTCH LIGAND INVOLVED IN THE MEDIATION OF NOTCH
RX SIGNALING. PLAYS AN ESSENTIAL ROLE DURING LIMB, CRANIOFACIAL, AND
RA DEVELOPMENT OF PERIPHERAL AND CENTRAL NERVOUS SYSTEMS.
RN [6]
RP TISSUE SPECIFICITY. FOUND TO BE HIGHEST IN FETAL THYMUS,
RX EPIDERMIS, FOREBRAIN, DORSAL ROOT GANGLIA AND INNER EAR. IN 2-WEEK-
RA OLD MICE IT WAS ABUNDANT IN HEART, LUNG, THYMUS, SKELETAL MUSCLE,
RN BRAIN AND TESTIS. EXPRESSION OVERLAPS PARTIALLY WITH NOTCH1
RX EXPRESSION.
RA [7]
RP DEVELOPMENTAL STAGE: AT 13 DPC IT IS FOUND IN PARAVERTEBRAL
RX VESSELS AND DORSAL ROOT GANGLIA. AT 14 DPC, IN OROPHARYNGEAL
RA EPITHELIUM, DEVELOPING THYMUS AND IN THE MUSCLES OF THE TONGUE. BY
RN 15 DPC, IN MANY TISSUES.
RX [8]
RP SIMILARITY: Contains 1 DSL domain.
RA [9]
RP SIMILARITY: Contains 16 EGF-like domains.
RX [10]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RA between the Swiss Institute of Bioinformatics and the EMBL Outstation-
RN the European Bioinformatics Institute. There are no restrictions on its
RX use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.isb-sib.ch/announce/
RX or send an email to license@sib-sib.ch).
CC EMBL: AF038572; AAF16411.1;
CC EMBL: AF010137; AAC14010.1;
CC EMBL: Y14495; CAA74835.1;
CC HSSP: P00743; IICF.
CC MGI: 1098870; Jag2.
CC GO: GO:0005887; C:integral to plasma membrane; IDA.
CC GO: GO:0008083; F:growth factor activity; ISS.
CC GO: GO:0005112; F:Notch binding activity; IPI.
CC GO: GO:0007049; F:cell cycle; ISS.
CC GO: GO:0030154; P:cell differentiation; NAS.
CC GO: GO:0001709; P:cell fate determination; NAS.
CC GO: GO:0007267; P:cell-cell signaling; NAS.

```

DR GO:0042492; P:gamma-delta T-cell differentiation; IMP.
 DR GO:0009912; P:hair cell fate commitment; IMP.
 DR GO:0007605; P:hearing; NAS.
 DR GO:0030326; P:limb morphogenesis; ISS.
 DR GO:0007219; P:N receptor signaling pathway; ISS.
 DR GO:0030334; P:regulation of cell migration; ISS.
 DR GO:0042127; P:regulation of cell proliferation; ISS.
 DR GO:0007283; P:spermatogenesis; ISS.
 DR GO:0045061; P:thymic T-cell selection; ISS.
 DR Interpro: IPR00152; Asx_hydroxyl.
 DR Interpro: IPR001774; DSL.
 DR Interpro: IPR000742; EGF_2.
 DR Interpro: IPR001881; EGF_Ca.
 DR Interpro: IPR001438; EGF_II.
 DR Interpro: IPR006209; EGF_III.
 DR Interpro: IPR001007; VWF_C.
 DR Pfam: PF01414; DSL; 1.
 DR Pfam: PF00008; EGF; 14.
 DR PRINTS: PR00010; EGFBIOD.
 DR SMART: SM00051; DSL; 1.
 DR SMART: SM00214; EGF_Ca; 9.
 DR SMART: SM00010; Asx_HYDROXYL; 10.
 DR PROSITE: PS00022; EGF_1; 16.
 DR PROSITE: PS01186; EGF_2; 11.
 DR PROSITE: PS01187; EGF_Ca; 7.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 23
 FT CHAIN 24 1247
 FT DOMAIN 24 1084
 FT TRANSMEM 1085 1107
 FT DOMAIN 1108 1247
 FT DOMAIN 178 240
 FT DOMAIN 241 274
 FT DOMAIN 275 305
 FT DOMAIN 307 345
 FT DOMAIN 347 383
 FT DOMAIN 385 421
 FT DOMAIN 423 459
 FT DOMAIN 461 496
 FT DOMAIN 498 534
 FT DOMAIN 536 572
 FT DOMAIN 574 634
 FT DOMAIN 636 672
 FT DOMAIN 674 710
 FT DOMAIN 712 748
 FT DOMAIN 751 787
 FT DOMAIN 789 825
 FT DOMAIN 827 863
 FT DOMAIN 865 901
 FT DISULFID 245 256
 FT DISULFID 258 269
 FT DISULFID 273 284
 FT DISULFID 287 298
 FT DISULFID 293 304
 FT DISULFID 304 315
 FT DISULFID 311 323
 FT DISULFID 317 333
 FT DISULFID 335 344
 FT DISULFID 344 356
 FT DISULFID 356 371
 FT DISULFID 373 382
 FT DISULFID 382 394
 FT DISULFID 394 409
 FT DISULFID 409 420
 FT DISULFID 420 438
 FT DISULFID 432 447
 FT DISULFID 449 458
 FT DISULFID 465 475
 FT DISULFID 469 484
 FT DISULFID 486 495
 FT DISULFID 502 513
 FT DISULFID 513 522

FT DISULFID 524 533 BY SIMILARITY.
 FT DISULFID 540 551 BY SIMILARITY.
 FT DISULFID 545 560 BY SIMILARITY.
 FT DISULFID 562 571 BY SIMILARITY.
 FT DISULFID 589 612 POTENTIAL.
 FT DISULFID 606 622 BY SIMILARITY.
 FT DISULFID 624 633 BY SIMILARITY.
 FT DISULFID 640 651 BY SIMILARITY.
 FT DISULFID 645 660 BY SIMILARITY.
 FT DISULFID 662 671 BY SIMILARITY.
 FT DISULFID 678 689 BY SIMILARITY.
 FT DISULFID 683 698 BY SIMILARITY.
 FT DISULFID 700 709 BY SIMILARITY.
 FT DISULFID 716 727 BY SIMILARITY.
 FT DISULFID 721 736 BY SIMILARITY.
 FT DISULFID 728 747 BY SIMILARITY.
 FT DISULFID 738 766 BY SIMILARITY.
 FT DISULFID 753 775 BY SIMILARITY.
 FT DISULFID 760 786 BY SIMILARITY.
 FT DISULFID 777 804 BY SIMILARITY.
 FT DISULFID 793 813 BY SIMILARITY.
 FT DISULFID 815 824 BY SIMILARITY.
 FT DISULFID 831 842 BY SIMILARITY.
 FT DISULFID 836 851 BY SIMILARITY.
 FT DISULFID 853 862 BY SIMILARITY.
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 570 570 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 619 619 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 302 302 L -> M (IN REF. 2).
 FT CONFLICT 461 461 N -> T (IN REF. 2).
 FT CONFLICT 469 478 COHGHGKRL -> VSAMGHGKLP (IN REF. 2).
 FT CONFLICT 492 492 G -> V (IN REF. 2).
 FT CONFLICT 546 546 L -> F (IN REF. 2).
 FT CONFLICT 549 549 A -> V (IN REF. 2).
 FT CONFLICT 735 738 RCAC -> PAR (IN REF. 3).
 FT CONFLICT 809 809 N -> H (IN REF. 2).
 FT CONFLICT 812 812 R -> A (IN REF. 2).
 SO SEQUENCE 1247 AA; 134726 MW; 1D80C8626FAFAEBC CRC64;
 Query Match 24.4%; Score 61.5; DB 1; Length 1247;
 Best Local Similarity 31.0%; Pred. No. 7.6;
 Matches 18; Conservative 1; Mismatches 16; Indels 23; Gaps 3;
 OY 4 ACTSYGCGRCGCASTHYLCRYL-----PKKMC-----ACVHCS 39
 DB 230 ACMDGMMGKRECAVCKOG-CNLLHGCTVPEBCRCSSYGKQKFCDECVYPGCVHGS 286
 RESULT 6
 PRTC_PIG STANDARD; PRT; 459 AA.
 AC 09GLP2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
 DE (Autoproteolysin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV).
 GN PROC.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=21121490; PubMed=11229814;
 RA Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,
 RA Kim H.K.W.;
 RT "Porcine factor V: cDNA cloning, gene mapping, three-dimensional protein modeling of membrane binding sites and comparative anatomy of

domains.";
 CC Cell. Mol. Life Sci. 56:148-159(2001).
 CC -1- FUNCTION: Protein C is a vitamin K-dependent serine protease that
 CC regulates blood coagulation by inactivating factors Va and VIIIa
 CC in the presence of calcium ions and phospholipids.
 CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
 CC and VIIIa.
 CC -1- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
 CC into a light chain and a heavy chain held together by a disulfide
 CC bond. The enzyme is then activated by thrombin, which cleaves a
 CC tetradecapeptide from the amino end of the heavy chain; this
 CC reaction, which occurs at the surface of endothelial cells, is
 CC strongly promoted by thrombomodulin.
 CC -1- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
 CC -1- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
 CC residues allows the modified protein to bind calcium.
 CC -1- MISCELLANEOUS: Calcium also binds, with stronger affinity to
 CC another site, beyond the Gla domain. This Gla-independent binding
 CC site is necessary for the recognition of the
 CC thrombin-thrombomodulin complex.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF191307; AAC28380.1; -
 DR HSSP; P04070; 1PCU.
 DR MEROPS; S01.218; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR006209; EGF_Like.
 DR InterPro: IPR002383; GLA_blood.
 DR InterPro: IPR006210; LEGF.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR InterPro: IPR000294; Vitk_dep_GLA.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00594; gla; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00001; GLABLOOD.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00069; GLA; 1.
 DR SMART; SM00020; TRYF_SPC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Blood coagulation; Glycoprotein; Serine protease;
 KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
 KW EGF-like domain; Repeat; Endothelial cell; Hydrolyase; Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 41
 FT CHAIN 42 459
 FT CHAIN 42 196
 FT CHAIN 199 459
 FT CHAIN 199 213
 FT PEPTIDE 199 213
 FT SITE 213 214
 FT DOMAIN 96 131
 FT DOMAIN 135 175
 EGF-LIKE 1.
 EGF-LIKE 2.

FT DOMAIN 214 459 SERINE PROTEASE.
 FT MOD_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID (BY
 FT MOD_RES 48 48 SIMILARITY).
 FT MOD_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID (BY
 FT MOD_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID (BY
 FT MOD_RES 57 57 SIMILARITY).
 FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID (BY
 FT MOD_RES 61 61 SIMILARITY).
 FT MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID (BY
 FT MOD_RES 67 67 SIMILARITY).
 FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID (BY
 FT MOD_RES 70 70 SIMILARITY).
 FT MOD_RES 112 112 HYDROXYLATION (BY SIMILARITY).
 FT ACT_SITE 255 255 CHARGE RELAY SYSTEM.
 FT ACT_SITE 301 301 CHARGE RELAY SYSTEM.
 FT ACT_SITE 400 400 CHARGE RELAY SYSTEM.
 FT DISULFID 58 63 BY SIMILARITY.
 FT DISULFID 91 110 BY SIMILARITY.
 FT DISULFID 100 105 BY SIMILARITY.
 FT DISULFID 104 119 BY SIMILARITY.
 FT DISULFID 121 130 BY SIMILARITY.
 FT DISULFID 139 150 BY SIMILARITY.
 FT DISULFID 146 159 BY SIMILARITY.
 FT DISULFID 161 174 BY SIMILARITY.
 FT DISULFID 182 321 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 240 256 BY SIMILARITY.
 FT DISULFID 371 385 BY SIMILARITY.
 FT DISULFID 396 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 459 AA; 51866 MW; 8541AAC14CC16D09 CRC64;
 Query Match 24.2%; Score 61; DB 1; Length 459;
 Best Local Similarity 31.0%; Pred. No. 3.6;
 Matches 13; Conservative 2; Mismatches 15; Indels 12; Gaps 2;
 QY 5 CTSWCGKFC-----GTASCTHYLCRVLPCKMCAC 35
 DB 121 CAQGWEGRFLCHEVRFNSCKSTENGCAHY-CLEEEGGRRCAC 161
 RESULT 7
 ID NX1L_NAJHH STANDARD; PRT; 71 AA.
 AC P25674;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Long neurotoxin 1 (Toxin CM-5).
 OS Naaja haje haje (Egyptian cobra).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Elapidae; Elapinae; Naaja.
 OX NCBI_TaxID=8642;
 RN [1]
 RP TISSUE-Venom.
 RC MEDLINE=79062487; PubMed=718974;
 RX Joubert F.J., Taljaard N.;
 RA "Purification, some properties and the primary structures of three
 RT reduced and S-carboxymethylated toxins (CM-5, CM-6 and CM-10a) from
 RT Naaja haje haje (Egyptian cobra) venom.";
 RL Biochim. Biophys. Acta 537:1-8(1978).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.

```

CC -1- MISCELLANEOUS: LD(50) IS 0.11 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: Belongs to the snake toxin family.
DR HSSP: P01391; 2CTX.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; toxin; 1.
DR ProDom: PD000206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Multigene family.
FT DISULFID 3 20 BY SIMILARITY.
FT DISULFID 14 41 BY SIMILARITY.
FT DISULFID 26 30 BY SIMILARITY.
FT DISULFID 45 56 BY SIMILARITY.
FT DISULFID 57 62 BY SIMILARITY.
SQ SEQUENCE 71 AA; 7821 MW; CFECS9A5A31802E3 CRC64;

Query Match 24.0%; Score 60.5; DB 1; Length 71;
Best Local Similarity 31.9%; Pred. No. 0.77;
Matches 15; Conservative 4; Mismatches 17; Indels 11; Gaps 4;

OY 3 HAC-TSYMGKFCG-----TASCTHYLCRVLHPG---KMACVHQS 39
DB 18 HVCYTKMCDNFCGMRGRVLDGCA-ATCPYKPGVDKCSIDNCH 63

RESULT 8
JAG2_HUMAN STANDARD: PRT: 1238 AA.
AC Q9Y219; Q9UE17; Q9UE99; Q9UNK8; Q9Y6P9; Q9Y6Q0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Jagged 2 precursor (Jagged2) (HJ2).
GN JAG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=97459705; PubMed=9315665;
RA Luo B., Aster J.C., Hassarjian R.P., Kuo F., Sklar J.;
RT "Isolation and functional analysis of a cDNA for human Jagged2, a gene
RL encoding a ligand for the Notch receptor.";
RN Mol. Cell. Biol. 17:6057-6067(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Fetal brain;
RX MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
RT "Human ligands of the Notch receptor.";
RN Am. J. Pathol. 154:785-794(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC TISSUE=Bone marrow;
RX MEDLINE=20130121; PubMed=10662552;
RA Deng Y., Medan A., Banta A.B., Friedman C., Trask B.J., Hood L.,
RA Li L.;
RT "Characterization, chromosomal localization, and the complete 30-kb
RL genomic sequence 63:133-138(2000).
RN [4]
RP SEQUENCE OF 17-1238 FROM N.A. (ISOFORM LONG).
RC TISSUE=Heart;
RX MEDLINE=98145947; PubMed=9486542;
RA Valsecchi C., Ghezzi C., Ballabio A., Ruggeri E.I.;
RT "JAGGED2: a putative Notch ligand expressed in the apical ectodermal
RL ridge and in sites of epithelial-mesenchymal interactions.";
Mech. Dev. 69:203-207(1997).
CC -1- FUNCTION: PUTATIVE NOTCH LIGAND INVOLVED IN THE MEDIATION OF NOTCH
CC SIGNALING. INVOLVED IN LIMB DEVELOPMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:

```

```

CC Event=Alternative splicing; Named isoforms=2;
CC Name=long;
CC IsoId=Q9Y219-1; Sequence=Displayed;
CC Name=short; Synonyms=HJAG2.del.E6;
CC IsoId=Q9Y219-2; Sequence=VSP_001395;
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND SKELETAL
CC MUSCLE AND TO A LESSER EXTENT IN PANCREAS. VERY LOW EXPRESSION IN
CC BRAIN, LUNG, LIVER AND KIDNEY.
CC -1- DISEASE: MAY BE ASSOCIATED TO USHER SYNDROME TYPE 1A (USH1A1)
CC WHICH DESCRIBES A CONGENITAL SENSOR DEAFNESS ASSOCIATED WITH
CC RETINITIS PIGMENTOSA AND FEEBLE-MINDEDNESS.
CC -1- SIMILARITY: Contains 16 EGF-like domains.
CC -1- SIMILARITY: Contains 1 DSL domain.
CC -1- SIMILARITY: Contains 1 VWF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: AF020201; AAB71189.1; -
DR EMBL: AF003521; AAB61285.1; -
DR EMBL: AF029778; AAB84215.1; -
DR EMBL: AF029779; AAB84216.1; -
DR EMBL: AF111170; AAD15562.1; -
DR EMBL: Y14330; CAA4706.1; -
DR HSSP: P00743; IACF.
DR GeneW: HGNC:6189; JAG2.
DR MIM: 602570; -
DR GO: GO:0005887; C: integral to plasma membrane; ISS.
DR GO: GO:0008083; F: growth factor activity; IDA.
DR GO: GO:0005112; F: Notch binding activity; IPL.
DR GO: GO:0007049; P: cell cycle; NAS.
DR GO: GO:0030154; P: cell differentiation; IDA.
DR GO: GO:0001709; P: cell fate determination; NAS.
DR GO: GO:0007267; P: cell-cell signaling; ISS.
DR GO: GO:0009912; P: hair cell fate commitment; ISS.
DR GO: GO:0007605; P: hearing; ISS.
DR GO: GO:0030326; P: limb morphogenesis; ISS.
DR GO: GO:0007219; P: N receptor signaling pathway; NAS.
DR GO: GO:0030334; P: regulation of cell migration; NAS.
DR GO: GO:0042127; P: regulation of cell proliferation; IDA.
DR GO: GO:0007283; P: spermatogenesis; IEP.
DR GO: GO:0030217; P: T-cell differentiation; IDA.
DR GO: GO:0045061; P: thymic T-cell selection; IDA.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_11.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PR00008; EGF_14.
DR PRINTS: PR00010; EGFBL00D.
DR SMART: SM00051; DSL; 1.
DR SMART: SM00179; EGF_CA; 9.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 10.
DR PROSITE: PS00022; EGF_1; 16.
DR PROSITE: PS01186; EGF_2; 12.
DR PROSITE: PS01187; EGF_CA; 7.
DR PROSITE: PS0184; VWF_C; 2; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
KW Repeat; Signal; Transmembrane; Alternative splicing.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1238 JAGGED 2.
FT DOMAIN 1081 1105 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1081 1105 POTENTIAL.
FT DOMAIN 1106 1238 CYTOPLASMIC (POTENTIAL).

```


05 Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 CC Xenopodinae; Xenopus.
 OX NCBI_TaxId=8353;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90040698; PubMed-2509712;
 RA Niefeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
 RA Poeling A., Knoechel W.;
 RT "Second-order repeats in Xenopus laevis finger proteins.";
 RL J. Mol. Biol. 208:639-659(1989).
 DR PIR: S06569; S06569.
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam: PF00096; zf_C2H2; 4.
 DR ProDom: PD000003; Znf_C2H2; 1.
 DR SMART: SM00355; Znf_C2H2; 4.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 4.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
 FT NON_TER 1 1
 FT ZN_FING 6 28 C2H2-TYPE.
 FT ZN_FING 34 56 C2H2-TYPE.
 FT ZN_FING 62 84 C2H2-TYPE.
 FT ZN_FING 89 111 C2H2-TYPE.
 FT NON_TER 111 111
 SQ SEQUENCE 111 AA; 12842 MW; 59DE9C06DDC76202 CRC64;
 Query Match 23.4%; Score 59; DB 1; Length 111;
 Best Local Similarity 42.9%; Pred. No. 1.7;
 Matches 18; Conservative 2; Mismatches 14; Indels 8; Gaps 4;
 OY 2 SHACTSYWCKRFGCTAS--CTHYLCRLVLPK-KMCACVHCSR 40
 DB 5 SFDCTE--CGKSFKRKSKTKTHFLC---HTGEKPPVCHGCK 41
 RESULT 11
 AT19_MOUSE STANDARD; PRT; 1210 AA.
 AC P59509;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE ADAMTS-19 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 with thrombospondin motifs 19) (ADAM-TS19) (ADAM-TS19).
 GN ADAMTS19.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6; TISSUE=Ovary;
 RX MEDLINE-22505168; PubMed-12617826;
 RA Menke D.B.; Page D.C.;
 RT "Sexually dimorphic gene expression in the developing mouse gonad.";
 RL Gene Expr. Patterns 2:359-367(2002).
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 matrix (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed predominantly in fetal ovary, low
 levels of expression is also detected in kidney, heart, skeletal
 muscle, lung and testis.
 CC -1- DEVELOPMENTAL STAGE: Expression is strongest in anterior and
 ventral regions of the ovary at 12.5 and 13.5 dpc before becoming
 more uniform.
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase (By
 similarity).
 CC -1- SIMILARITY: Belongs to peptidase family M12b.
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -1- SIMILARITY: Contains 5 TSP type-1 domains.
 CC -1- CAUTION: By homology with the human sequence, it is uncertain

CC whether Met-1 or Met-5 is the initiator.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as their content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AY135183; AA10155.1; -.
 DR MGD: MGI:2442875; Adamts19.
 DR SMART: SM00608; ACR; 1.
 DR SMART: SM00209; TSP1; 5.
 DR PROSITE: PS50215; ADAM_MPRO; 1.
 DR PROSITE: PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE: PS50214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE: PS50092; TSP1; 5.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KM Repeat; Extracellular matrix.
 FT SIGNAL 1 30
 FT PROPEP 31 319
 FT CHAIN 320 1210
 FT DOMAIN 320 548
 FT DOMAIN 549 636
 FT DOMAIN 637 689
 FT DOMAIN 690 793
 FT DOMAIN 794 917
 FT DOMAIN 918 978
 FT DOMAIN 979 1040
 FT DOMAIN 1042 1086
 FT DOMAIN 1090 1147
 FT DOMAIN 1167 1198
 FT SITE 297 297
 FT METAL 485 485
 FT ACT_SITE 486 486
 FT METAL 489 489
 FT METAL 495 495
 FT CARBOHYD 54 54
 FT CARBOHYD 263 263
 FT CARBOHYD 800 800
 FT CARBOHYD 910 910
 FT CARBOHYD 931 931
 FT CARBOHYD 952 952
 FT CARBOHYD 1012 1012
 SQ SEQUENCE 1210 AA; 134560 MW; 0AB812ABAB4BB7A2 CRC64;
 Query Match 23.4%; Score 59; DB 1; Length 1210;
 Best Local Similarity 36.4%; Pred. No. 14;
 Matches 16; Conservative 1; Mismatches 15; Indels 12; Gaps 3;
 OY 5 CTSYWGKRGCTASCTHYL-----CRLVLPKMKACVHCSR 40
 DB 591 CTGLWC-KVEGEAEKRTKLPMDGTC---DPKWKCKGEGCTR 630
 RESULT 12
 AGI_HORVU STANDARD; PRT; 212 AA.
 ID AGI_HORVU
 AC P15312;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Root-specific lectin precursor.
 OS Hordeum vulgare (Barley).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 CC Triticeae; Hordeum.
 OX NCBI_TaxId=4513;
 RN (1)
 RP SEQUENCE FROM N.A.

RA Lerner D.R., Raskhel N.V.;
 RT "Cloning and characterization of root-specific barley lectin."
 RL Plant Physiol. 91:124-129(1989).
 CC -1- FUNCTION: CARBOHYDRATE BINDING.
 CC -1- SIMILARITY: Contains 4 chitin-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M29280; AAA32969.1; -.
 DR PIR: T05936; T05936.
 DR HSSP: P10969; 1MGT.
 DR InterPro: IPR001002; Chitin_binding_1.
 DR Pfam: PF00187; Chitin_bind.1; 4.
 DR PRINTS: PR00451; CHITINBINDNG.
 DR SMART: SM00270; ChEBD1.4.
 DR PROSITE: PS00026; CHITIN_BINDING; 4.
 KW Lectin; Repeat; Chitin-binding; Glycoprotein; Signal;
 FT Pyroglutamate carboxylic acid.
 FT SIGNAL 1 26
 FT CHAIN 27 212
 FT DOMAIN 27 69 ROOT-SPECIFIC LECTIN.
 FT DOMAIN 70 112 CHITIN-BINDING 1.
 FT DOMAIN 113 155 CHITIN-BINDING 2.
 FT DOMAIN 156 197 CHITIN-BINDING 3.
 FT DOMAIN 27 27 CHITIN-BINDING 4.
 FT MOD_RES 27 27 PYRROLIDONE CARBOXYLIC ACID
 FT (BY SIMILARITY).
 FT DISULFID 29 44 BY SIMILARITY.
 FT DISULFID 38 50 BY SIMILARITY.
 FT DISULFID 43 57 BY SIMILARITY.
 FT DISULFID 61 66 BY SIMILARITY.
 FT DISULFID 72 87 BY SIMILARITY.
 FT DISULFID 81 93 BY SIMILARITY.
 FT DISULFID 86 100 BY SIMILARITY.
 FT DISULFID 104 109 BY SIMILARITY.
 FT DISULFID 115 130 BY SIMILARITY.
 FT DISULFID 124 136 BY SIMILARITY.
 FT DISULFID 129 143 BY SIMILARITY.
 FT DISULFID 147 152 BY SIMILARITY.
 FT DISULFID 158 173 BY SIMILARITY.
 FT DISULFID 167 179 BY SIMILARITY.
 FT DISULFID 172 186 BY SIMILARITY.
 FT DISULFID 190 195 BY SIMILARITY.
 FT CARBOHYD 206
 FT N-LINKED (GLCNAC...)
 SO SEQUENCE 212 AA; 21209 MW; 8D948245D6B625A5 CRC64;
 Query Match 23.2%; Score 58.5; DB 1; Length 212;
 Best local Similarity 32.6%; Pred. No. 3.5;
 Matches 14; Conservative 3; Mismatches 17; Indels 9; Gaps 2;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Dueserhoeft A., Floeth M., Fritze C., Heuss-Netzel D.,
 RA Hilbert H., Moestl D.,
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Andre B., Iraqi Houssein I., Urrestarazu L.A., Vissers S.,
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO DROSOPHILA SHUTTLE CRAFT PROTEIN (STC) AND HUMAN
 CC NFX1.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z71299; CAA95885.1; -.
 DR PIR: S62935; S62935.
 DR SCD: S0004968; PAP1.
 DR InterPro: IPR001374; R3H.
 DR InterPro: IPR000967; ZnF_NFX1.
 DR InterPro: IPR001841; ZnF_Ring.
 DR Pfam: PF01424; R3H; 1.
 DR Pfam: PF01422; zf-NFX1; 5.
 DR SMART: SM00393; R3H; 1.
 DR SMART: SM0038; ZnF_NFX; 7.
 DR PROSITE: PS00518; ZF_RING_1; FALSE-NEG.
 DR PROSITE: PS50089; ZF_RING_2; 1.
 KW Hypothetical protein; zinc-finger; Repeat.
 FT ZN_FING 68 118
 FT DOMAIN 150 644 7 X APPROXIMATE REPEATS, CYS-RICH.
 FT REPEAT 150 185 1.
 FT REPEAT 206 251 2.
 FT REPEAT 273 330 3.
 FT REPEAT 352 390 4.
 FT REPEAT 458 497 5.
 FT REPEAT 575 610 6.
 FT REPEAT 611 644 7.
 SO SEQUENCE 965 AA; 108494 MW; 121C57BB07C6FA9D CRC64;
 Query Match 23.2%; Score 58.5; DB 1; Length 965;
 Best local Similarity 26.3%; Pred. No. 13;
 Matches 10; Conservative 6; Mismatches 15; Indels 7; Gaps 1;

OY 5 CTSY-----WCGKFCGTASC-THYLCRVLHPKMCACV 38
 ID YNC3_YEAST STANDARD; PRT; 965 AA.
 AC P53971;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical 108.5 kDa protein in UME3-HDA1 intergenic region.
 DE YNM023C OR N2812.
 GN Saccharomyces cerevisiae (Baker's yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Ascomycotales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Jagl Brare
 RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Oda T., Chandrasekharappa S.C.,
 RL Isolation, characterization and expression analysis of zebrafish
 RT jagged genes.";
 ID JAG1_BRARE STANDARD; PRT; 1242 AA.
 AC O9057;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Jagged 1 precursor (Jagged1).
 GN JAG1.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RT Oda T., Chandrasekharappa S.C.,
 RL Isolation, characterization and expression analysis of zebrafish
 RT jagged genes.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 CC -1- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling (By similarity). Seems to be involved
 CC in cell-fate decisions.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: Contains 17 EGF-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF229448; AAL08213.1; -
 DR ZFIN; ZDB-GENE-011128-2; Jag1.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001774; DSL.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR006209; EGF_Like.
 DR InterPro: IPR002049; Laminin_EGF.
 DR Pfam: PF01414; DSL; 1.
 DR Pfam: PF00008; EGF; 14.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR SMART; SM00051; DSL; 1.
 DR SMART; SM00179; EGF_Ca; 9.
 DR SMART; SM00214; VMC; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 9.
 DR PROSITE: PS00022; EGF_1; 16.
 DR PROSITE: PS01186; EGF_2; 12.
 DR PROSITE: PS01187; EGF_Ca; 8.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
 DR Repeat; Transmembrane; Signal.
 KW SIGNAL.
 FT CHAIN 1 28
 FT DOMAIN 29 1242 JAGGED 1.
 FT TRANSMEM 1071 1095 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1096 1242 POTENTIAL.
 FT DOMAIN 168 230 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 231 264 DSL.
 FT DOMAIN 262 295 EGF-Like 1.
 FT DOMAIN 297 335 EGF-Like 2.
 FT DOMAIN 337 373 EGF-Like 3.
 FT DOMAIN 375 411 EGF-Like 4.
 FT DOMAIN 413 449 EGF-Like 5.
 FT DOMAIN 451 486 EGF-Like 6.
 FT DOMAIN 488 524 EGF-Like 7.
 FT DOMAIN 526 562 EGF-Like 8.
 FT DOMAIN 575 630 EGF-Like 9.
 FT DOMAIN 632 668 EGF-Like 10.
 FT DOMAIN 670 706 EGF-Like 11.
 FT DOMAIN 708 744 EGF-Like 12.
 FT DOMAIN 747 783 EGF-Like 13.
 FT DOMAIN 785 821 EGF-Like 14.
 FT DOMAIN 823 859 EGF-Like 15.
 FT DOMAIN 859 917 EGF-Like 16.
 FT DOMAIN 917 959 EGF-Like 17.
 FT DOMAIN 959 1000 EGF-Like 18.
 FT DISULFID 235 246 BY SIMILARITY.
 FT DISULFID 239 252 BY SIMILARITY.
 FT DISULFID 254 263 BY SIMILARITY.
 FT DISULFID 266 277 BY SIMILARITY.
 FT DISULFID 272 283 BY SIMILARITY.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 301 313 BY SIMILARITY.
 FT DISULFID 307 323 BY SIMILARITY.
 FT DISULFID 325 334 BY SIMILARITY.
 FT DISULFID 341 352 BY SIMILARITY.
 FT DISULFID 346 361 BY SIMILARITY.
 FT DISULFID 363 372 BY SIMILARITY.
 FT DISULFID 379 390 BY SIMILARITY.

FT DISULFID 384 399 BY SIMILARITY.
 FT DISULFID 401 410 BY SIMILARITY.
 FT DISULFID 417 428 BY SIMILARITY.
 FT DISULFID 432 437 BY SIMILARITY.
 FT DISULFID 439 448 BY SIMILARITY.
 FT DISULFID 455 465 BY SIMILARITY.
 FT DISULFID 469 474 BY SIMILARITY.
 FT DISULFID 476 485 BY SIMILARITY.
 FT DISULFID 492 503 BY SIMILARITY.
 FT DISULFID 497 512 BY SIMILARITY.
 FT DISULFID 514 523 BY SIMILARITY.
 FT DISULFID 530 541 BY SIMILARITY.
 FT DISULFID 535 550 BY SIMILARITY.
 FT DISULFID 552 561 BY SIMILARITY.
 FT DISULFID 566 647 BY SIMILARITY.
 FT DISULFID 641 656 BY SIMILARITY.
 FT DISULFID 658 667 BY SIMILARITY.
 FT DISULFID 674 685 BY SIMILARITY.
 FT DISULFID 679 694 BY SIMILARITY.
 FT DISULFID 696 705 BY SIMILARITY.
 FT DISULFID 712 723 BY SIMILARITY.
 FT DISULFID 717 732 BY SIMILARITY.
 FT DISULFID 734 743 BY SIMILARITY.
 FT DISULFID 751 762 BY SIMILARITY.
 FT DISULFID 756 771 BY SIMILARITY.
 FT DISULFID 773 782 BY SIMILARITY.
 FT DISULFID 789 800 BY SIMILARITY.
 FT DISULFID 794 809 BY SIMILARITY.
 FT DISULFID 811 820 BY SIMILARITY.
 FT DISULFID 827 838 BY SIMILARITY.
 FT DISULFID 832 847 BY SIMILARITY.
 FT DISULFID 849 858 BY SIMILARITY.
 FT CARBOHYD 141 141 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 385 385 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 748 748 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 991 991 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1046 1046 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 1242 AA; 135343 MW; 59557A4B7F052860 CRC64;
 Query Match 23.2%; Score 58.5; DB 1; Length 1242;
 Best Local Similarity 35.9%; Pred. No. 17;
 Matches 14; Conservative 3; Mismatches 13; Indels 9; Gaps 3;
 QY 5 CTSTWCGKFCGTASC-----THYLGRVLHPGKMCACVH 37
 DB 221 CLEGGWGEPCNTAICKGCSIEHGSCKV--PGR-CRCLY 256
 RESULT 15
 ID FAT_DROME STANDARD; PRT; 5147 AA.
 AC P33450; Q9Y0X5;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cadherin-related tumor suppressor precursor (Fat protein).
 GN FT OR CG3352.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92069752; PubMed=1959133;
 RA Mahoney P.A., Weber U., Onofrechuk P., Bliessmann H., Bryant P.J.,
 RA Goodman C.S.;
 RT "The fat tumor suppressor gene in Drosophila encodes a novel member
 of the cadherin gene superfamily.";
 RL Cell 67:853-868(1991).


```
FT DISULFID 4334 4350 BY SIMILARITY.
FT DISULFID 4352 4361 BY SIMILARITY.
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 605 605 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1155 1155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1367 1367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1458 1458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1831 1831 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1880 1880 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2080 2080 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2171 2171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2247 2247 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2290 2290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2437 2437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2581 2581 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2799 2799 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2920 2920 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2946 2946 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2967 2967 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3167 3167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3303 3303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3386 3386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3389 3389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3525 3525 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3852 3852 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3865 3865 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3905 3905 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 23.28; Score 58.5; DB 1; Length 5147;
Best Local Similarity 32.1%; Pred. No. 60;
Matches 18; Conservative 3; Mismatches 14; Indels 21; Gaps 4;
```

```
OY 6 TSYWC-----GKCGTAS-----CTH-YLCRVLHFGKMKACV-----HCSR 40
DB 4074 SSYFCLRPGRFGNOCESVSDSCRPNCLHGLCVSLKPGYKNCCTPGRYGRHCR 4129
```

Search completed: September 3, 2003, 10:59:32
Job time : 14.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:43:26 ; Search time 60 Seconds

(without alignments)
105.818 Million cell updates/sec

Title: US-10-030-231-5

Perfect score: 199
Sequence: 1 HXHXCTSYCKKFCGTAXCTRYXCRLHXKXCXCHCSR 40

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	88.9	96	22	ABR76984
2	175	87.9	40	22	ABR76986
3	175	87.9	96	22	ABR76985
4	66.5	33.4	479	22	ABR95891
5	62	31.2	3907	23	ABG70822
6	62	31.2	3907	24	ABG74190
7	60	30.2	7285	24	ABU38280
8	57	28.6	101	22	AAG73959
9	57	28.6	109	17	AA84086

10	56	28.1	1435	24	ABU07908
11	55	27.6	430	20	AA31745
12	54	27.1	71	24	ABR98360
13	54	27.1	4440	24	ABU07907
14	54	27.1	7339	24	AA016358
15	53.5	26.9	55	20	AAV40052
16	53.5	26.9	708	22	ABG20107
17	53	26.6	102	22	AAO3473
18	53	26.6	274	22	ABR58794
19	52.5	26.4	212	15	AA45357
20	52.5	26.4	212	21	AA19645
21	52.5	26.4	882	22	ABR70358
22	52	26.1	93	22	AAO0301
23	52	26.1	603	22	ABR5123
24	51.5	25.9	662	22	ABR63237
25	51.5	25.9	1329	23	AAU91279
26	51.5	25.9	2211	24	ABG72959
27	51	25.6	282	22	ABR59436
28	51	25.6	1172	23	ABR72334
29	51	25.6	1743	24	ABU07906
30	50.5	25.4	114	23	ABP00604
31	50.5	25.4	863	22	ABR64979
32	50	25.1	116	22	AAH88752
33	49.5	24.9	1300	24	ABU07905
34	49	24.6	379	13	AA25565
35	49	24.6	379	22	AAE05920
36	49	24.6	379	22	ABR09201
37	48.5	24.4	75	21	AAV90579
38	48.5	24.4	122	21	AAH42599
39	48.5	24.4	277	23	ABR97515
40	48.5	24.4	513	22	ABR73600
41	48.5	24.4	518	23	ABR05388
42	48.5	24.4	540	22	ABG25308
43	48.5	24.4	606	22	ABR57274
44	48.5	24.4	925	23	AAO14246
45	48.5	24.4	1044	23	ABP64908

ALIGNMENTS

RESULT 1	ABR76984	standard; Protein; 96 AA.
ID	ABR76984;	
AC	22-JUL-2002 (first entry)	
DT	Antimicrobial peptide Myticline a.	
XX	Myticline a: mollusc; microbial disease; antimicrobial; antibacterial; fungicidal.	
KW	Mytilus gallioprovincialis.	
XX		
OS	Mytilus gallioprovincialis.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..20
FT	Peptide	/label= Signal_peptide
FT	Peptide	21..60
FT	Protein	/note= "This region is specifically claimed in Claim 3"
FT	Protein	21..96
FT	Protein	/label= Mature_protein
XX		
PN	FR2796072-A1.	
XX		
PD	12-JAN-2001.	
XX		
PF	08-JUL-1999;	99FR-0008858.
XX		
PR	08-JUL-1999;	99FR-0008858.
XX		
PA	(CNRS) CNRS CENT NAT RECH SCI.	

PA	(IFREMER INST FR RECH EXPL MER.
XX	
PI	Roach P, Maita G, Hubert F, Noel T;
XX	
DR	WPI: 2001-149782/16.
DR	N-PSDB; ABL58046.
XX	
PS	New antimicrobial peptides myticines obtainable from a bivalve mollusc,
PT	especially Mytilus galloprovincialis are useful for treatment and
PT	prevention of microbial disease -
XX	
PS	Disclosure; Page 13; 18pp; French.
XX	
CC	The present invention relates to new antimicrobial peptides (ABB76984 and
CC	ABB76985), named myticines, obtainable from a bivalve mollusc (Mytilus
CC	galloprovincialis). The peptides have a molecular weight of about 4.5 kD,
CC	an isoelectric point of about 8.7 and comprise 8 cysteine residues. The
CC	present sequence is myticine a. The peptides have antibacterial and
CC	fungicidal activity and can be used to prepare anti-infective medicaments
CC	and to prevent and treat microbial diseases in various sectors, e.g.
CC	health, agriculture, aquaculture and animal husbandry.
XX	
SQ	Sequence 96 AA;
QY	Query Match 88.9%; Score 177; DB 22; Length 96;
Db	Best Local Similarity 70.0%; Pred. No. 4.6e-14;
	Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
	1 HXHACTSYXCXKFCGTANCTYYXRXLHXGXKXCXCHCSR 40
	1 11111111111111111111111111111111111111
	21 HSHACTSYWCGKFCGTASTCYHLCLRVLHPGKMCACVHCSSR 60
RESULT 2	
ABB76986	
ID	ABB76986 standard; Peptide; 40 AA.
XX	
AC	ABB76986;
XX	
DT	22-JUL-2002. (first entry)
XX	
DE	Antimicrobial peptide myticine consensus sequence.
XX	
KW	Myticine; mollusc; microbial disease; antimicrobial; antibacterial;
KW	fungicidal.
XX	
OS	Mytilus galloprovincialis.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 2 /label= Pro, Ser
FT	Misc-difference 4 /label= Val, Ala
FT	Misc-difference 9 /label= Try, Trp
FT	Misc-difference 11 /label= Ser, Gly
FT	Misc-difference 18 /label= Ser, Gly
FT	Misc-difference 21 /label= Ser, Gly
FT	Misc-difference 23 /label= Arg, His
FT	Misc-difference 26 /label= Gly, Leu
FT	Misc-difference 29 /label= Asn, Val
FT	Misc-difference 32 /label= Arg, Pro
FT	Misc-difference 34 /label= Leu, Met
FT	Misc-difference 36 /label= Phe, Ala
FT	Misc-difference /label= Leu, His

XX FR2796072-A1.
XX
XX 12-JAN-2001.
XX
XX 08-JUL-1999; 99FR-0008858.
XX
XX 08-JUL-1999; 99FR-0008858.
XX
XX (CNRS) CNRS CENT NAT RECH SCI.
XX PA (IFREMER) IFREMER INST FR RECH EXPL MER.
XX PI Roach P, Milla G, Hubert F, Noel T;
XX WPI: 2001-149782/16.
DR
XX New antimicrobial peptides myticines obtainable from a bivalve mollusc,
PT especially Mytilus galloprovincialis are useful for treatment and
PT prevention of microbial disease -
PS Claim 2; Page 15; 18pp; French.

XX The present invention relates to new antimicrobial peptides (ABB76984 and
CC ABB76985), named myticines, obtainable from a bivalve mollusc (Mytilus
CC galloprovincialis). The peptides have a molecular weight of about 4.5 kD,
CC an isoelectric point of about 8.7 and comprise 8 cysteine residues. The
CC present sequence is a consensus sequence for the myticine peptides. The
CC peptides have antibacterial and fungicidal activity and can be used to
CC prepare anti-infective medicaments and to prevent and treat microbial
CC diseases in various sectors, e.g. health, agriculture, aquaculture and
CC animal husbandry.

SQ Sequence 40 AA:

Query Match 87.9%; Score 175; DB 22; Length 40;
Best Local Similarity 100.0%; Pred. No. 3; 9e-14;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 HXHXTSYXCXKFCGTAKCTYYXCRXLHGXKCXCHCSR 40
|||||
1 HHXCTSYXCXKFCGTAKCTYYXCRXLHGXKCXCHCSR 40

DB 1 HHXCTSYXCXKFCGTAKCTYYXCRXLHGXKCXCHCSR 40

RESULT 3
ABB76985
ID ABB76985 standard; Protein; 96 AA.
XX ABB76985;
AC
XX 22-JUL-2002 (first entry)
DT
XX Antimicrobial peptide Myticine b.
DE
XX Antimicrobial peptide Myticine b.
XX Myticine b; mollusc; microbial disease; antimicrobial; antibacterial;
KM fungicidal.
XX
OS Mytilus galloprovincialis.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label= Signal_peptide 21..60
FT /note= "This region is specifically claimed in Claim 3"
FT Protein 21..96
FT /label= Mature_protein
XX
PN FR2796072-A1.
PD 12-JAN-2001.
PF 08-JUL-1999; 99FR-0008858.
PR 08-JUL-1999; 99FR-0008858.

XX (CNRS) CNRS CENT NAT RECH SCI.
PA (IFREMER) IFREMER INST FR RECH EXPL MER.
XX
XX Roch P, Mittle G, Hubert F, Noel T;
XX WPI: 2001-149782/16.
DR N-PSDB; ABL58047.
XX
XX New antimicrobial peptides myticines obtainable from a bivalve mollusc,
PT especially Mytilus galloprovincialis are useful for treatment and
PT prevention of microbial disease -
XX
XX Disclosure: Page 14; 18pp; French.
XX
XX The present invention relates to new antimicrobial peptides (AB76984 and
CC AB76985), named myticines, obtainable from a bivalve mollusc (Mytilus
CC galloprovincialis). The peptides have a molecular weight of about 4.5 kD,
CC an isoelectric point of about 8.7 and comprise 8 cysteine residues. The
CC present sequence is myticine b. The peptides have antibacterial and
CC fungicidal activity and can be used to prepare anti-infective medicaments
CC and to prevent and treat microbial diseases in various sectors, e.g.
CC health, agriculture, aquaculture and animal husbandry.
XX
SQ Sequence 96 AA;

Query Match 87.9%; Score 175; DB 22; Length 96;
Best Local Similarity 70.0%; Pred. No. 8e-14; Mismatches 0; Gaps 0;
Matches 28; Conservative 0; Indels 12;

OY 1 HXHXCTSYXCKFCGTACTYXXCRXLHXGKXCXCHCSR 40
| | | | | | | | | | | | | | | | | | | | | |
DB 21 HPHVCTSYXCKFCGTACTYXXCRXLHXGKXCXCHCSR 60

RESULT 4
AAB95891
ID AAB95891 standard; Protein: 479 AA.

AC AAB95891;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:19019.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isegaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX

PS Claim 8; SEQ ID 19019; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to a
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX

SQ Sequence 479 AA;

Query Match 33.4%; Score 66.5; DB 22; Length 479;
Best Local Similarity 42.1%; Pred. No. 2.4; Mismatches 15; Gaps 4;
Matches 16; Conservative 2; Indels 15;

OY 7 SYXCKFCGTA-XCTYXXCR--XLHXGKX-CXCHCSR 40
: | | | | | | | | | | | | | | | | | | | | | |
DB 121 AYECKT-CGKAFKCPSYLCRHVTHSGKKFCCKCKGK 157

RESULT 5
ABG70822
ID ABG70822 standard; Protein: 3907 AA.

AC ABG70822;

DT 16-DEC-2002 (first entry)

DE Mouse myocardin associated protein #1.

XX Mouse; myocardin; myocardial infarction; cardiomyocyte;

KW post-mitotic cell; differentiation; therapeutic; gene therapy;

KW heart disease; cardiomyopathy; mortality; heart failure; hypertension.

OS Mus sp.

PN WO200260946-A2.

PD 08-AUG-2002.

PR 21-DEC-2001; 2001WO-US50606.

PR 21-DEC-2000; 2000US-257716P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Olson EN, Wang D;

XX WPI: 2002-732693/79.

XX New myocardin polypeptides and polynucleotides, useful for respecifying
PT non-cardiac cells; stimulating cardiac tissue regeneration, and for
PT treating cardiovascular disorders, such as myocardial infarction and
PT hypertension -
XX
XX Disclosure: Page 146-159; 175pp; English.

CC The invention discloses an isolated polynucleotide encoding a myocardin
CC polypeptide. Myocardial infarction results in the loss of cardiomyocytes,
CC which are post-mitotic cells and generally do not regenerate after birth.
CC Transplanting foetal cardiomyocytes has limitations so identifying new
CC regulators of cardiomyocyte growth and differentiation is an important
CC goal in the search for therapeutics to treat myocardial tissue damage.
CC The polynucleotides, polypeptides and methods of the invention can be
CC used to modulate the phenotype of a non-cardiomyocyte cell, to generate a
CC or more phenotypic functions of a cardiomyocyte cell, to generate a
CC cardiomyocyte, which comprises introducing into a cardiac fibroblast the
CC myocardin polynucleotide and a promoter which is active in the
CC fibroblast, stimulate cardiac tissue regeneration which comprises
CC inhibiting the function of myocardin in a post-mitotic cardiomyocyte and
CC for screening for a modulator of myocardin expression. The nucleic acid
CC can also be used in gene therapy to treat a heart disease, including a
CC cardiomyopathy, comprising administering a polynucleotide encoding a
CC myocardin peptide or protein or an antisense nucleic acid. The
CC polynucleotide and polypeptide can also be used for decreasing mortality
CC in a subject with heart failure, comprising inhibiting the function of
CC myocardin in post-mitotic cardiomyocytes, increasing the level of
CC myocardin in fibroblasts to generate cardiomyocytes, inhibiting the
CC function of myocardin in post-mitotic cardiomyocytes or increasing the
CC level of myocardin in fibroblasts to generate cardiomyocytes in the
CC subject. The compositions and methods of the present invention are useful
CC for respecifying non-cardiac cells, stimulating cardiac tissue
CC regeneration, for treating cardiovascular disorders, such as myocardial
CC infarction and hypertension, and for screening compounds for various
CC abilities to interact and/or affect myocardin expression or function. The
CC sequence presented is the mouse myocardin associated protein, #1.
CC Note: This sequence is presented as a 3-letter coded protein sequence in
CC the specification, but when changed into the single letter code, forms a
CC DNA sequence identical to that given in ABS55224. It, therefore, appears
CC to be the DNA sequence encoding human myocardin 2.

XX Sequence 3907 AA;

Query Match 31.2%; Score 62; DB 23; Length 3907;
Best Local Similarity 34.3%; Pred. No. 48;
Matches 12; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 5 CTSYXCKRFGCTAXCTYXCRKLMHGKXCXCHCS 39
DB 1455 CACCTCTGACGCGAGCTGACCTGCAGGCTCGCCA 1489

RESULT 6
ABG74190
ID ABG74190 standard; Protein; 3907 AA.

XX ABG74190;

DT 29-APR-2003 (first entry)

DE Mouse myocardin associated protein.

XX Mouse; myocardin; cardiomyocyte growth; ANF induction;
KM cardiomyocyte differentiation; sarcomere assembly induction; ANF;
KW cardiac hypertrophy; atrial natriuretic factor; fibroblast modulation;
KW non-cardiomyocyte cell modulation; cardiomyocyte; heart disease;
KM cardiomyopathy; myocardial infarction; hypertension; gene therapy;
KW cardiac-specific transcription factor.

OS Mus musculus.

PN US2002164735-A1.

PD 07-NOV-2002.

PF 21-DEC-2001; 2001US-0029217.

PR 21-DEC-2000; 2000US-257761P.

PA (OLSON) OLSON E. N.

PA (WANG/) WANG D.

XX Olson EN, Wang D;

DR WPI: 2003-247258/24.

PT Novel isolated polynucleotide encoding human or murine myocardin 1
PT polypeptide, useful for modulating phenotype of non-cardiomyocyte cell
PT e.g., fibroblast, to include phenotypic functions of cardiomyocyte cell
PS Disclosure; Page 51-61; 104pp; English.

XX The invention relates to an isolated polynucleotide encoding myocardin
XX polypeptide. The effects of myocardin in growth and/or all
XX differentiation of cardiomyocytes was assessed by overexpressing
XX myocardin in cardiomyocytes using adenoviral delivering system. The
XX results showed that overexpression of myocardin in neonatal
XX cardiomyocytes induced assembly of sarcomeres and expression of atrial
XX natriuretic factor, ANF, markers of cardiac hypertrophy. An expression
XX cassette containing the polynucleotide operably linked to a regulatory
XX sequence is useful for modulating the phenotype of a non-cardiomyocyte
XX cell e.g. fibroblast to include one or more phenotypic functions of a
XX cardiomyocyte cell. The expression cassette is useful for generating a
XX cardiomyocyte which involves introducing into a cardiac fibroblast the
XX expression cassette comprising the polynucleotide and a promoter active
XX in the fibroblast, where the promoter directs the expression of the
XX polynucleotide. The expression cassette further comprises a second
XX polynucleotide encoding GATA4, under the control of a second promoter
XX active in a cardiac fibroblast. The expression cassette further comprises
XX a polyadenylation site and an immunological marker. An expression
XX cassette comprising the polynucleotide encoding myocardin protein or
XX peptide and a promoter operable in eukaryotic cells is useful for
XX treating a heart disease, including cardiomyopathy, such as myocardial
XX infarction or hypertension. The present sequence represents the amino
XX acid sequence of the mouse myocardin associated protein.
XX Note: The protein sequence presented is not disclosed in the
XX specification but is shown in the sequence listing.

XX Sequence 3907 AA;

Query Match 31.2%; Score 62; DB 24; Length 3907;
Best Local Similarity 34.3%; Pred. No. 48;
Matches 12; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 5 CTSYXCKRFGCTAXCTYXCRKLMHGKXCXCHCS 39
DB 1455 CACCTCTGACGCGAGCTGACCTGCAGGCTCGCCA 1489

RESULT 7
ABJ38280
ID ABJ38280 standard; Protein; 7285 AA.

XX ABJ38280;

DT 12-JUN-2003 (first entry)

DE PAMG21-RANK-Fc vector protein SEQ ID No 28.

XX TALL-1-binding protein; TALL-1; B-cell-mediated autoimmune disease;
KM systemic lupus erythematosus; B-cell-mediated cancer; lymphoma;
KW inflammation; rheumatoid arthritis; acute pancreatitis; atherosclerosis;
KW Alzheimer's disease; asthma; cachexia; cirrhosis; diabetes; osteoporosis;
KW glomerulonephritis; hashimoto's thyroiditis; ischaemic injury; psoriasis;
KW multiple myeloma; multiple sclerosis; Parkinson's disease; vasculitis;
KW gene therapy; ds.

OS Unidentified.

PN W0200292620-A2.

PD 21-NOV-2002.

XX 13-MAY-2002; 2002WO-US15273.
XX
XX
PR 11-MAY-2001; 2001US-290196P.
XX
XX (AMGE-) AMGEN INC.
XX
XX Min H, Hsu H;
XX
DR WPI; 2003-156719/15.
XX
XX New TALL-1-binding polypeptide, useful for modulating the activity of
PT TALL-1 and in treating, preventing or diagnosing a B-cell-mediated
PT autoimmune diseases, cancers or lymphomas -
XX
XX
PS Disclosure; Fig 4; 236pp; English.
XX
XX The invention relates to a novel TALL-1-binding polypeptide comprising a
CC defined sequence in the specification. The composition is useful in
CC modulating the activity of TALL-1, and in treating, preventing,
CC ameliorating, diagnosing or prognosing a B-cell-mediated autoimmune
CC disease (e.g. systemic lupus erythematosus) or B-cell-mediated cancer or
CC lymphoma. The composition may also be used in treating inflammations
CC (e.g. rheumatoid arthritis), acute pancreatitis, Alzheimer's disease,
CC asthma, atherosclerosis, cachexia, cirrhosis, diabetes,
CC glomerulonephritis, Hashimoto's thyroiditis, ischaemic injury, multiple
CC myeloma, multiple sclerosis, osteoporosis, Parkinson's disease, psoriasis
CC and vasculitis. Disorders may be treated with the novel composition using
CC gene therapy. This polynucleotide sequence represents a TALL-1 related
CC DNA sequence of the invention.
XX
SQ Sequence 7285 AA;
XX
Query Match 30.2%; Score 60; DB 24; Length 7285;
Best Local Similarity 32.3%; Pred. No. 1.4e+02;
Matches 10; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
XX
OY 5 CTSYXCKFCGTAXCTYXXCRXLHGKXCXC 35
DB 6009 CCCAGCACCCGTTGCAGCTCACACAGGACAC 6039
XX
RESULT 8
AAG73959
ID AAG73959 standard; protein; 101 AA.
XX
AC AAG73959;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:4723.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX
XX Homo sapiens.
XX
XX MO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX WPI; 2001-235357/24.
DR N-PSDB; AAH33390.
XX

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX
PS Claim 11; Page 6525-6526; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 101 AA;
XX
Query Match 28.6%; Score 57; DB 22; Length 101;
Best Local Similarity 29.4%; Pred. No. 8.9;
Matches 10; Conservative 2; Mismatches 22; Indels 0; Gaps 0;
XX
OY 5 CTSYXCKFCGTAXCTYXXCRXLHGKXCXCXC 38
DB 37 CTGFCVCVCVCACACVLCFLHSLPLSLYHC 70
XX
RESULT 9
AAR84086
ID AAR84086 standard; protein; 109 AA.
XX
XX AAR84086;
XX
XX 28-NOV-1996 (first entry)
XX
DE T-lymphocyte stimulatory protein.
XX
XX E. maxima; T-lymphocyte stimulatory protein; Eimeria; protozoan;
KW coccidiosis; chicken; vaccine; poultry; probe.
XX
XX Eimeria maxima.
XX
XX AU9531720-A.
XX
XX 28-MAR-1996.
XX
XX 15-SEP-1995; 95AU-0031720.
PF 16-SEP-1994; 94EP-0202676.
PR
XX
XX (ALKU) AKZO NOBEL NV.
XX
XX Bumstead JM, Dunn PFD, Tomley FM, Vermeulen AN;
PI
XX WPI; 1996-210114/22.
DR N-PSDB; AAT14351.
XX
XX DNA encoding Eimeria T-lymphocyte stimulatory protein - used in
PT vaccines to protect poultry against coccidiosis, and to develop
PT prods. for diagnosis of Eimeria infection
XX
XX Claim 1; Page 46; 59pp; English.
PS
XX This sequence represents E. maxima T-lymphocyte stimulatory protein.
XX

CC Eimeria protozoans are the cause of coccidiosis in chickens. The DNA
CC encoding this sequence may be attached to a suitable promoter and used
CC in a recombinant vector in the production of a vaccine for the protection
CC of poultry against coccidiosis. Fragments of this sequence may also
CC be used as probes to detect Eimeria-related nucleic acid in tissue.
CC Due to poor print quality in the specification, this sequence is a
CC "best guess" based on the corresponding DNA sequence.

SQ Sequence 109 AA;

Query Match	28.6%;	Score 57;	DB 17;	Length 109;
Best Local Similarity	30.6%;	Pred. No. 9.5;		
Matches	11;	Conservative	4;	Mismatches 17;
			Indels	4;
			Gaps	1

```

QY      5 CTSTYXCKKFCGTACTYXXCRRLHMGKXXCXKXCSR 40
         |::|||::|::|::|
Db      40 CSTYCCSTFCCCKCCCKCFCNCFNRFC---CSR 71

```

RESULT 10
ABU07908
ID ABU07908 standard; Protein; 1435 AA.

AC ABU07908;

DT 20-MAY-2003 (first entry)

DE Novel human secreted and transmembrane protein #5.

KM Human, secreted protein; transmembrane protein; cytosolic;
 KM gene therapy; TNF-agonist-Alpha; chondrocyte stimulator; tumour
 KM adrenal tumour; lung tumour; colon tumour; breast tumour;
 KM prostate tumour; rectal tumour; cervical tumour; liver tumour.

05 Homo sapiens.

PN US2003036157-A1.

PD 20-FEB-2003.

02-JUL-2002; PF

PR 16-SEP-1998;

PR 01-DEC-1998;

PR 14-MAY-1999;

PR 01-SEP-1999

PR 01-DEC-1999;

PR 30-DEC-1999

PR 18-FEB-2000;

PR 22-FEB-2000;

PR 01-MAR-2000;

PR 15-MAR-2000;

PR 17-MAY-2000;

PR 30-MAY-2000

PR 28-JUL-2000;

PR 08-NOV-2000;

PR 20-DEC-2000;

PR , 01-JUN-2001

PR	20-JUN-2001	2001NO-US19692
PR	29-JUN-2001	2001NO-US21066
PR	09-JUL-2001	2001NO-US21363
PR	29-AUG-2001	2001NO-US227099
PR	18-SEP-1997	97US-US928263
PR	18-SEP-1997	97US-US928266
PR	17-OCT-1997	97US-US622520
PR	21-OCT-1997	97US-US63466
PR	24-OCT-1997	97US-US63120
PR	24-OCT-1997	97US-US63121P
PR	28-OCT-1997	97US-US635410P
PR	28-OCT-1997	97US-US63541P
PR	28-OCT-1997	97US-US63544P
PR	28-OCT-1997	97US-US63564P
PR	29-OCT-1997	97US-US63734P
PR	31-OCT-1997	97US-US63870
PR	31-OCT-1997	97US-US64103P
PR	13-NOV-1997	97US-US65311P
PR	21-NOV-1997	97US-US66120P
PR	24-NOV-1997	97US-US66466P
PR	24-NOV-1997	97US-US67712P
PR	11-DEC-1997	97US-US69335P
PR	12-DEC-1997	97US-US69425P
PR	17-DEC-1997	97US-US69870
PR	18-DEC-1997	97US-US68017P
PR	10-MAR-1998	98US-US77430P
PR	11-MAR-1998	98US-US77632P
PR	11-MAR-1998	98US-US77666P
PR	20-MAR-1998	98US-US78866P
PR	20-MAR-1998	98US-US78939P
PR	27-MAR-1998	98US-US79664P
PR	27-MAR-1998	98US-US79786P
PR	31-MAR-1998	98US-US80107P
PR	31-MAR-1998	98US-US80143P
PR	01-APR-1998	98US-US80337P
PR	01-APR-1998	98US-US80333P
PR	08-APR-1998	98US-US81049P
PR	08-APR-1998	98US-US81070P
PR	09-APR-1998	98US-US81135P
PR	15-APR-1998	98US-US81838P
PR	21-APR-1998	98US-US82568P
PR	21-APR-1998	98US-US82569P
PR	22-APR-1998	98US-US82704P
PR	22-APR-1998	98US-US82737P
PR	28-APR-1998	98US-US83345P
PR	29-APR-1998	98US-US83452P
PR	29-APR-1998	98US-US83466P
PR	29-APR-1998	98US-US83499P
PR	29-APR-1998	98US-US83559P
PR	05-MAY-1998	98US-US84356P
PR	06-MAY-1998	98US-US84414P
PR	07-MAY-1998	98US-US84659P
PR	07-MAY-1998	98US-US84660P
PR	15-MAY-1998	98US-US85362P
PR	15-MAY-1998	98US-US85360P
PR	15-MAY-1998	98US-US85362P
PR	18-MAY-1998	98US-US86300P
PR	22-MAY-1998	98US-US86332P
PR	22-MAY-1998	98US-US86466P
PR	28-MAY-1998	98US-US87059P
PR	28-MAY-1998	98US-US87208P
PR	02-JUN-1998	98US-US87609P
PR	02-JUN-1998	98US-US87759P
PR	03-JUN-1998	98US-US87827P
PR	04-JUN-1998	98US-US88028P
PR	04-JUN-1998	98US-US88028P
PR	04-JUN-1998	98US-US88033P
PR	04-JUN-1998	98US-US88326P
PR	05-JUN-1998	98US-US88167P
PR	05-JUN-1998	98US-US88202P

PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088555P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-091632P.
PR 04-AUG-1998; 98US-094006P.
PR 10-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
Query Match 28.1%; Score 56; DB 24; Length 1435;
Best Local Similarity 34.3%; Pred. No. 1.1e+02;
Matches 12; Conservative 1; Mismatches 16; Indels 6; Gaps 2;
Oy 5 CTSYXCKRFGCTAXCTYXCRXLHGKXCCXCHCS 39
DB 994 CTG--CCACCGTCGCG---CATTCGACGACCACT 1022
RESULT 11
ID AAY31745 standard; Protein; 430 AA.
XX AAY31745;
AC AAY31745;
XX 22-NOV-1999 (first entry)
DT 22-NOV-1999 (first entry)
XX
DE Mycobacterium tuberculosis specific DNA-encoded polypeptide.
XX Tuberculosis; infection; diagnosis; DNA probe.
OS Mycobacterium tuberculosis.
XX
PH key Location/Qualifiers
FT Misc-difference 4 /note= "encoded by TGA"
FT Misc-difference 6 /note= "encoded by TGA"
FT Misc-difference 20 /note= "encoded by TGA"
FT Misc-difference 29 /note= "encoded by TGA"
FT Misc-difference 29 /note= "encoded by TGA"
FT Misc-difference 54 /note= "encoded by TGA"
FT Misc-difference 54 /note= "encoded by TGA"
FT Misc-difference 64 /note= "encoded by TGA"
FT Misc-difference 69 /note= "encoded by TGA"
FT Misc-difference 69 /note= "encoded by TGA"
FT Misc-difference 89 /note= "encoded by TGA"
FT Misc-difference 89 /note= "encoded by TGA"
FT Misc-difference 99 /note= "encoded by TGA"
FT Misc-difference 114 /note= "encoded by TGA"
FT Misc-difference 114 /note= "encoded by TGA"
FT Misc-difference 119 /note= "encoded by TGA"
FT Misc-difference 119 /note= "encoded by TGA"
FT Misc-difference 129 /note= "encoded by TGA"
FT Misc-difference 129 /note= "encoded by TGA"
FT Misc-difference 159 /note= "encoded by TGA"
FT Misc-difference 159 /note= "encoded by TGA"
FT Misc-difference 169 /note= "encoded by TGA"
FT Misc-difference 169 /note= "encoded by TGA"
FT Misc-difference 182 /note= "encoded by TGA"
FT Misc-difference 182 /note= "encoded by TAG"
FT Misc-difference 185 /note= "encoded by TGA"
FT Misc-difference 185 /note= "encoded by TGA"
FT Misc-difference 219 /note= "encoded by TGA"
FT Misc-difference 219 /note= "encoded by TGA"
FT Misc-difference 259 /note= "encoded by TGA"
FT Misc-difference 259 /note= "encoded by TGA"
FT Misc-difference 269 /note= "encoded by TGA"
FT Misc-difference 269 /note= "encoded by TGA"
FT Misc-difference 291 /note= "encoded by TGA"
FT Misc-difference 291 /note= "encoded by TGA"
FT Misc-difference 323 /note= "encoded by TGA"
FT Misc-difference 323 /note= "encoded by TGA"
FT Misc-difference 339 /note= "encoded by TGA"
FT Misc-difference 339 /note= "encoded by TGA"
FT Misc-difference 349 /note= "encoded by TGA"
FT Misc-difference 349 /note= "encoded by TGA"

FT	Misc-difference	356	/note=	"encoded by TGA"
FT	Misc-difference	366	/note=	"encoded by TGA"
FT	Misc-difference	410	/note=	"encoded by TGA"
FT	Misc-difference	421	/note=	"encoded by TGA"
FT	Misc-difference	421	/note=	"encoded by TGA"
XX	EP45462-AL.			
PN	29-SEP-1999.			
PD	25-MAR-1998;	98EP-0302287.		
XX	25-MAR-1998;	98EP-0302287.		
XX	25-MAR-1998;	98EP-0302287.		
XX	(COUL) CSIR COUNCIL SCI IND RES.			
PA	Kumar D, Srivastava BS, Srivastava R;			
XX	WPI: 1999-530042/45.			
XX	N-PSDB; AAX87940.			
XX	New nucleic acid molecules, useful for detecting and identifying			
XX	Mycobacterium tuberculosis			
XX	Disclosure: Page 15-27; 42pp; English.			
XX	The present sequence represents an amino acid sequence deduced			
CC	from the Mycobacterium tuberculosis specific DNA fragment provided			
CC	in AAX87940. This DNA fragment comprises a StuI-StuI fragment of			
CC	M. tuberculosis genomic DNA and contains an insertion sequence-like			
CC	element and repetitive sequences. The DNA fragment is useful as a			
CC	probe, especially for detecting or identifying M. tuberculosis in			
CC	clinical isolates and body fluids e.g. sputum, cerebrospinal fluid,			
CC	pleural fluid, urine, gastric lavage, bronchial lavage, pericardial			
CC	or lymph node aspirate (all claimed). It is also useful for			
CC	restriction fragment length polymorphism analysis of M. tuberculosis			
CC	isolates (claimed). The probe provides rapid and specific diagnosis			
CC	of tuberculosis and M. tuberculosis infection.			
XX				
SO	Sequence	430 AA;		
Query Match	27.6%;	Score 55;	DB 20;	Length 430;
Best Local Similarity	35.1%;	Pred. No. 52;		
Matches 13;	Conservative	2;	Mismatches	20;
			Indels	2;
			Gaps	2;
Oy	5 CTSYCKKF-CGTAKCTYXXCRXLHGXKXCXCHSR	40		
Db	146 CPELRRCRRRCPRXRQGRWCPN-XNGRCRCRPPSSR	181		
RESULT 12				
AC	ABB99360			
XX	ABB99360; standard; Protein; 71 AA.			
XX	29-JAN-2003 (first entry)			
XX	Amino acid sequence of human T23490 gene.			
DE	Human; T23490; benign prostatic hyperplasia.			
XX	Human; T23490; benign prostatic hyperplasia.			
OS	Homo sapiens.			
XX	Key	Location/Qualifiers		
XX	Domain	11..34		
FT	Modified-site	17		
FT	Modified-site	27		
FT	Modified-site	27		

FT		/note= "N-myristoylation site"
FT	Modified-site	42
FT		/note= "N-myristoylation site"
FT	Modified-site	52
FT		/note= "N-myristoylation site"
FT	Modified-site	54
FT		/note= "potential casein kinase II phosphorylation site"
FT	Modified-site	63
FT		/note= "N-myristoylation site"
XX		
XX	WO200277162-A2.	
XX		
PD	03-OCT-2002.	
XX		
PF	14-MAR-2002; 2002MO-US07728.	
XX		
PR	22-MAR-2001; 2001US-277634P.	
XX		
PA	(GENE-) GENE LOGIC INC.	
PA	(NISB) JAPAN TOBACCO INC.	
XX		
PI	Kulkarni P, Getzenberg RH, Munger WE;	
XX		
DR	WPI; 2003-029922/02.	
DR	N-PSDB; ABV72464.	
XX		
PT	New T23490 nucleic acid molecule useful as a diagnostic agent or marker	
PT	to diagnose or monitor benign prostatic hyperplasia or its progression,	
PT	or to assess prostatic function -	
XX		
XS	Claim 19; Page 49-50; 50pp; English.	
XX		
CC	The present sequence is encoded by an open reading frame of a gene	
CC	designated T23490, which is associated with benign prostatic hyperplasia.	
CC	The T23490 nucleic acid molecule and the encoded protein are useful as	
CC	diagnostic agents or markers to detect or monitor benign prostatic	
CC	hyperplasia or its progression. These can also serve as a target for	
CC	agents that can be used to modulate gene expression or the activity of	
CC	the protein. The proteins may also be used for the screening of synthetic	
CC	small molecules and combinatorial or naturally-occurring compound	
CC	libraries to discover novel therapeutics to regulate prostatic function.	
XX		
SQ	Sequence 71 AA;	
	Query Match 27.1%; Score 54; DB 24; Length 71.	
	Best Local Similarity 25.5%; Pred. No. 15;	
	Matches 12; Conservative 3; Mismatches 20; Indels 12; Gaps 2	
OY	3 HXCTSYXCKKF-----CGTAXCTYXYXCRLHNGKXCX-CXH 37	
	: :	
Dd	14 HLCGSHLCOAYLRGSHLCRSHLCGYVLGGSYLCRSPDAGLTCAECTY 60	
	:	
RESULT 13		
ID	ABU07907	
XX	ABU07907 standard; Protein; 4440 AA.	
AC	ABU07907;	
XX		
DT	20-MAY-2003 (first entry)	
XX		
DE	Novel human secreted and transmembrane protein #4.	
XX		
KW	Human; secreted protein; transmembrane protein; cytosolic;	
KW	gene therapy; TNF-agonist-Alpha; chondrocyte stimulator; tumour;	
KW	adrenal tumour; lung tumour; colon tumour; breast tumour;	
KW	prostate tumour; rectal tumour; cervical tumour; liver tumour.	
XX		
OS	Homo sapiens.	
XX		
PJ	US2003036157-A1.	
XX		
PD	20-FEB-2003.	

XX 02-JUL-2002; 2002US-0188769.
PF 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-052963P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.

PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 09-JUN-1998; 98US-088657P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-0900246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090546P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.

PR 26-JUN-1998; 98US-090863P.
 PR 26-JUN-1998; 98US-091010P.
 PR 01-JUL-1998; 98US-091359P.
 PR 01-JUL-1998; 98US-091544P.
 PR 02-JUL-1998; 98US-091478P.
 PR 02-JUL-1998; 98US-091486P.
 PR 02-JUL-1998; 98US-091626P.
 PR 02-JUL-1998; 98US-091628P.
 PR 02-JUL-1998; 98US-091632P.
 PR 02-JUL-1998; 98US-094006P.
 PR 04-AUG-1998; 98US-095282P.
 PR 10-AUG-1998; 98US-095912P.
 PR 10-AUG-1998; 98US-096012P.
 PR 17-AUG-1998; 98US-096757P.
 PR 17-AUG-1998; 98US-096766P.
 PR 17-AUG-1998; 98US-096867P.
 PR 17-AUG-1998; 98US-096891P.
 PR 17-AUG-1998; 98US-096897P.
 PR 18-AUG-1998; 98US-096949P.
 PR 18-AUG-1998; 98US-096959P.
 PR 18-AUG-1998; 98US-097022P.
 PR 26-AUG-1998; 98US-097952P.
 PR 26-AUG-1998; 98US-097954P.
 PR 26-AUG-1998; 98US-097955P.
 PR 26-AUG-1998; 98US-097971P.
 PR 26-AUG-1998; 98US-097974P.
 PR 26-AUG-1998; 98US-098014P.
 PR 01-SEP-1998; 98US-098166P.
 PR 01-SEP-1998; 98US-098723P.
 PR 02-SEP-1998; 98US-098803P.
 PR 02-SEP-1998; 98US-098821P.
 PR 02-SEP-1998; 98US-098843P.
 PR 09-SEP-1998; 98US-099602P.
 PR 10-SEP-1998; 98US-099741P.
 PR 10-SEP-1998; 98US-099754P.
 PR 10-SEP-1998; 98US-099763P.

Query Match 27.1%; Score 54; DB 24; Length 4440;
 Best Local Similarity 28.6%; Pred. No. 4.8e+02;
 Matches 10; Conservative 2; Mismatches 15; Indels 8; Gaps 1;

OY 5 CTSYXCKKFCGTAXCTYXCRXLHXKCKXCHCS 39
 DB 3315 CCTATCCTTCGTCAATGTTCTCTTCA 3341

RESULT 14

AA016358 standard; Protein; 7339 AA.

ID AA016358 standard; Protein; 7339 AA.
 AC AA016358;
 XX
 DT 03-APR-2003 (first entry)
 XX
 DE Human translocated promoter region (TPR) protein, SEQ ID No. 6.
 XX
 KW Human; p53 pathway modulating agent; p53-associated disorder;
 KW translocated promoter region; TPR; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200299050-A2.
 XX
 PD 12-DEC-2002.
 XX
 PE 03-JUN-2002; 2002WO-US17425.
 XX
 PR 05-JUN-2001; 2001US-296076P.
 PR 10-OCT-2001; 2001US-328605P.
 PR 15-FEB-2002; 2002US-357253P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX

PI Friedman L, Plozman GD, Belvin M, Francis-Jang H;
 XX
 DR WPI; 2003-167335/16.
 XX
 PT Identifying candidate p53 pathway-modulating agents, useful as
 PT therapeutic targets for disorders associated with defective p53
 PT function, comprises screening for agents that modulate the activity of
 PT translocated promoter region polypeptides -
 XX
 PS Claim 13; Page 53-84; 94pp; English.

CC The invention comprises a method for identifying a candidate p53 pathway
 CC modulating agent. The method involves providing an assay system
 CC comprising a purified translocated promoter region (TPR) protein or
 CC nucleic acid. The method of the invention is useful for identifying
 CC candidate p53 pathway modulating agents for use as therapeutic targets
 CC for disorders associated with defective p53 function (e.g. cancer). The
 CC TPR proteins and nucleic acids are useful for identifying agents that
 CC modulate TPR function. The present amino acid sequence represents a human
 CC TPR protein.
 CC NOTE: The present amino acid sequence appears to be the DNA sequence from
 CC SEQ ID No 5 (AAL51476) shown as a protein.

SO Sequence 7339 AA;

Query Match 27.1%; Score 54; DB 24; Length 7339;
 Best Local Similarity 25.7%; Pred. No. 7.3e+02;
 Matches 9; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

OY 5 CTSYXCKKFCGTAXCTYXCRXLHXKCKXCHCS 39
 DB 5004 CTGACACGACAGACAGATCATCTGAAACAACT 5038

RESULT 15
 ID AA40052 standard; Peptide; 55 AA.

AC AA40052;

DT 18-NOV-1999 (first entry)

DE Peptide sequence derived from a human secreted protein.

KW Secreted protein; gene therapy; cancer; tumor; fetal deficiency;
 KW neurodegenerative disorder; developmental abnormality; blood disorder;
 KW immune system disease; autoimmune disease; leukemia; inflammation;
 KW allergy; Alzheimer's disease; cognitive disorder; schizophrenia;
 KW obesity; osteoporosis; arthritis; infection; AIDS; diabetes; asthma;
 KW connective tissue disorder; transplant rejection; sepsis; acne;
 KW psoriasis; cardiovascular disorder; reproductive disorder;
 KW food additive; food preservative; storage capability.

XX Homo sapiens.

PN WO9943693-A1.

PD 02-SEP-1999.

PF 24-FEB-1999; 99WO-US03939.

PR 26-FEB-1998; 98US-0076051.

PR 26-FEB-1998; 98US-0076052.

PR 26-FEB-1998; 98US-0076053.

PR 26-FEB-1998; 98US-0076054.

PR 26-FEB-1998; 98US-0076057.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Olsen HS, Florence K, Brewer LA, Ebner R, Ruben SM, Rosen CA;
 PI Duan RD;
 DR WPI; 1999-550857/46.

XX New human genes and the secreted polypeptides they encode, useful for
 PT diagnosis and treatment of e.g. cancers, neurological disorders, immune
 PT diseases, inflammation or blood disorders

XX
 PS Disclosure: Page 44; 246pp; English.

CC AAY40001-92 are derived from human secreted proteins. The
 CC polynucleotides and their corresponding secreted polypeptides are useful
 CC for preventing, treating or ameliorating medical conditions, e.g. by
 CC protein or gene therapy. Pathological conditions can also be diagnosed by
 CC determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the polynucleotide. Specific
 CC uses include developing products for the diagnosis or treatment of
 CC cancer, tumors, neurodegenerative disorders, developmental abnormalities
 CC and fetal deficiencies, blood disorders, sepsis, diseases of the immune
 CC system, autoimmune diseases, inflammation, allergies, Alzheimer's and
 CC cognitive disorders, schizophrenia, obesity, osteoporosis, arthritis,
 CC infections, AIDS, connective tissue disorders, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC and reproductive disorders. The polypeptides or polynucleotides can
 CC also be used as food additives or preservatives, such as to increase
 CC or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors or other nutritional
 CC components.

XX
 SQ Sequence 55 AA:

Query Match 26.9%; Score 53.5; DB 20; Length 55;
 Best Local Similarity 38.7%; Pred. No. 14;
 Matches 12; Conservative 1; Mismatches 15; Indels 3; Gaps 1;

QY 13 FCGTAXC---TXYGCRXLHXGKXCXCHCSR 40
 |||||:| | | | | | | | | |
 Db 5 FCGTSSCLGTAVRCRAPAPWWSVRCPCR 35

Search completed: September 3, 2003, 10:58:56
 Job time : 62 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:54:26 ; Search time 60.5 Seconds
(without alignments)
170.613 Million cell updates/sec

Title: US-10-030-231-5
Perfect score: 199
Sequence: 1 HXHCTSYXCKFCGTACXTYXCRLHXGKXCXCHCSR 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66.5	33.4	479	4	Q96JVI
2	59.5	29.9	461	4	Q9HD74
3	56.5	28.4	648	11	Q8K167
4	55	27.6	195	5	Q8T3C5
5	54.5	27.4	625	11	Q8BIN6
6	54	27.1	4823	13	Q93321
7	53.5	26.9	595	4	Q8N1R9
8	53.5	26.9	595	4	Q8N1R9
9	53	26.6	771	11	Q8C020
10	52.5	26.4	585	5	Q9O0E2
11	52.5	26.4	882	5	Q9V3J9
12	52	26.1	603	5	Q9V525
13	51.5	25.9	163	12	Q8VB89
14	51.5	25.9	662	5	Q9VAV6
15	51	25.6	282	5	Q9W230
16	50.5	25.4	464	11	Q9JURI

17	50.5	25.4	464	11	Q9JUR3	Q9JUR3 mus musculu
18	50.5	25.4	514	5	Q9VRV3	Q9VRV3 drosophila
19	50.5	25.4	788	5	Q96444	Q96444 biophalar
20	50	25.1	85	5	Q9NL72	Q9NL72 caenorhabd
21	50	25.1	144	16	Q8FY49	Q8FY49 bruceella su
22	50	25.1	226	10	Q9FR5	Q9FR5 oryza sativ
23	50	25.1	344	5	Q900K3	Q900K3 plasmodi
24	50	25.1	783	5	P92163	P92163 strongyloce
25	50	25.1	1137	16	Q8CZU0	Q8CZU0 versinia pe
26	49.5	24.9	313	11	Q8B1K3	Q8B1K3 mus musculu
27	49.5	24.9	597	11	Q8CFE9	Q8CFE9 mus musculu
28	49	24.6	352	5	Q902C9	Q902C9 caenorhabd
29	49	24.6	379	11	Q9WTM9	Q9WTM9 ratus norv
30	48.5	24.4	197	6	Q28584	Q28584 ovis arie
31	48.5	24.4	276	4	Q9NPP0	Q9NPP0 homo sapien
32	48.5	24.4	411	13	P70046	P70046 xenopus lae
33	48.5	24.4	502	4	Q8N7M2	Q8N7M2 homo sapien
34	48.5	24.4	502	11	Q8R5B3	Q8R5B3 mus musculu
35	48.5	24.4	516	11	Q8C668	Q8C668 mus musculu
36	48.5	24.4	528	11	Q8R2M7	Q8R2M7 mus musculu
37	48.5	24.4	606	11	Q64247	Q64247 mus musculu
38	48.5	24.4	612	11	Q8BP18	Q8BP18 mus musculu
39	48.5	24.4	654	6	Q9SK52	Q9SK52 macaca fasc
40	48.5	24.4	1044	4	Q8WX49	Q8WX49 homo sapien
41	48.5	24.4	1044	4	Q9NXC6	Q9NXC6 homo sapien
42	48.5	24.4	1338	4	Q9BYF0	Q9BYF0 homo sapien
43	48	24.1	120	11	Q9CV14	Q9CV14 mus musculu
44	48	24.1	167	11	Q9D122	Q9D122 mus musculu
45	48	24.1	478	5	Q8SW72	Q8SW72 encephallo

ALIGNMENTS

RESULT 1
ID Q96JVI PRELIMINARY: PRT: 479 AA.
AC Q96JVI: TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 23, Last annotation update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ14959.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isoaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Komura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.,
RT "NEO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AK027865; BAB55416.1; -
DR InterPro: IPR007087; Znf_C2H2.
DR InterPro: IPR007086; Znf_C2H2_sub.
DR InterPro: IPR006025; Znf_MTPetide.
DR Pfam: PF00096; zfc_C2H2; 6.
DR PRINTS: PR00048; ZINC_FINGER.
DR SMART: SM00355; ZNF_C2H2; 6.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 5.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Metal-binding; Nuclear protein; zinc;
KW zinc-finger.
SQ SEQUENCE 479 AA: 55354 MW: B180AEK85F132319 CRC64;
Query Match 33.4%; Score 66.5; DB 4; Length 479;
Best Local Similarity 42.1%; Pred. No. 0.0097;

	Matches	16; Conservative	2;	Mismatches	15;	Indels	5;	Gaps	4;
Oy	7	SYXCKKFCGTA-XCTYYXCR--XLHXGX-CXCHCSR	40	:					:
		: :							
Db	121	AVECTK-CGRAFKCPSYLCRHEVTHSGKKPCECKQCGK	157						

RESULT 2	
Q9HD74	
ID Q9HD74	PRELIMINARY; . PRT; 461 AA

DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Zinc finger protein SBZ73.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute-
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Hom-
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang W., Wan T., Gao X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases
 DR EMBL; AF242519; AAF969600.1; -
 DR HSSB; P07248; ZADR.
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam: PF01352; KRAB; 1.
 DR Pfam: PF00096; Zf_C2H2; 9.
 DR ProDom: PD000003; Znf_C2H2; 7.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 7.
 DR PROSITE; PSS0805; KRAB; 1.
 DR PROSITE; PSS0028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PSS0157; ZINC_FINGER_C2H2_2; 11.
 DR Metal-binding; Zinc; zinc-finger.
 SQ SEQUENCE 461 AA; 5401 MW; 7C71BC3A6F50ADB CRC64;

Query Match	29.98;	Score 59.5;	DB 4;	Length 461;
Best Local Similarity	38.98;	Pred. No. 0.12;		
Matches 14;	Conservative 5;	Mismatches 12;	Indels 5;	Gaps 4

QY 8 YXCKAFECGTA--XCTYYX-CRLHXG-KKXCXCHES 39
| | | | | : : : : : | | | :
Db 212 YQCK-CGKAINECSGFTDCKRIHNGEHCKCEEN 240

RESULT 3		
08K167		
ID 08K167	PRELIMINARY;	PRT; 648 AA

DT .01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Similar to zinc finger protein 208.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC028265; AAH28265.1; -
DR InterPro: IPR007087; znf_C2H2.
DR Pfam: PF000096; zf_C2H2; 16.
DR SMART: SMO0355; znf_C2H2; 20.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 19.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 22.
DR Metal-Binding; Zinc; zinc-finger.
SQ SEQUENCE 648 AA; 74210 MW; 3985D3369D46A481 CRC64;

Query Match	28.4%	Score 56.5;	DB 11;	Length 648;
Best Local Similarity	35.1%;	Pred. No. 0.49;		
Matches 13;	Conservative	5;	Mismatches 14;	Indels 5;
				Gaps 3;

```

0y      8 YXCKKFGTAQXT---YXCRXLHXGXK-CXCHCSR 40
      | | | | | | | : : | | | | | :
Db      453 YNC-KVGNAFCTPSSLIQHKRIHTGELPCKDYCGK 480

```

RESULT 4	
Q8T3C5	
ID Q8T3C5	PRELIMINARY;
	PRT; 195 AA

DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE T22H6.7 protein (ABF-6).
GN T22H6.7 OR ABF-6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Rhabditida; Rhabditidae.
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI-TaxID=6239;

RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL submitted (AUG-1995) to the EMBL/GenBank/DBJ databases

RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; Pubmed=9851916;

RT "Genome sequence of the nematode *C.elegans*: A platform for
RT investigating biology."; <https://doi.org/10.1126/science.282.2012-2018>.
RL Science 282:2012-2018(1998).

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA zhang H., Kato Y.;

RT "abf-6, a member of ASABF-type antimicrobial peptide family in
RT *Caenorhabditis elegans*."; [PMID:12125205](#)
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z50757; C122638.1; -
DR EMBL; AB086410; BAC44993.1; -
DR WormPep; T22H6.7; CE30355.
SQ SEQUENCE 195 AA; 21158 MW; 11DA9F544ABBD63F CRC64

Query Match	27.68;	Score 55;	DB 5;	Length 195;
Best Similarity	28.68;	Pred. No. 0.31;		
Matches 10; Conservative	3;	Mismatches 20;	Indels 2;	Gaps 1

QY 6 TSYXCKPFGTAXCTYYXCRLHXGR--XCXCHC 38
 | | : | : | : | : | : | : |
Db 34 TDPICTSMCKVRFCSGSCRSVMGSGDPTCEECSC 66

RESULT 5	
Q8BIN6	
ID 08ATN6	PRELIMINARY:
	PRT: 625 AA

RC 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to MS2F5-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Mus.
OX NCBI-TaxId=10090;
RN [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX MEDLINE=2354683; PubMed=12466851;
RA The FANOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK038563; BAC30046.1;
 SQ SEQUENCE 625 AA; 72214 MW; 9747A245D04A82BE CRC64;

Query Match 27.4%; Score 54.5; DB 11; Length 625;
 Best Local Similarity 35.1%; Pred. No. 0.98;
 Matches 13; Conservative 4; Mismatches 15; Indels 5; Gaps 3;

OY 8 YXCKKFCGTACTXXC---RXLHXGX-CXCKHCSR 40
 DB 165 YNC-EVCGKAFCTSLFSKXKTHHGNKPCCEVCGK 200

RESULT 6
 OY 093321 PRELIMINARY; PRT; 4823 AA.
 AC 093321;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE All-1 related protein.
 GN ALR.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;
 CC Tetraodontidae; Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99177347; PubMed=10077531;
 RA Gellner K., Brenner S.;
 RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
 rubripes.";
 RL Genome Res. 9:251-258(1999).
 CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
 DR EMBL: AF056116; AAC34383.1; -
 DR InterPro: IPR003889; FYRICH.C.
 DR InterPro: IPR003888; FYRICH.N.
 DR InterPro: IPR000910; HMG_12_box.
 DR InterPro: IPR003616; PostSET.
 DR InterPro: IPR001214; SET.
 DR InterPro: IPR001965; ZnF_PHD.
 DR InterPro: IPR001841; ZnF_Ring.
 DR Pfam: PF00628; PHD; 5.
 DR Pfam: PF00856; SET; 1.
 DR SMART: SM00542; FYRC; 1.
 DR SMART: SM00541; FYRN; 1.
 DR SMART: SM00398; HMG; 1.
 DR SMART: SM00249; PHD; 8.
 DR SMART: SM00508; PostSET; 1.
 DR SMART: SM00184; RING; 4.
 DR SMART: SM00317; SET; 1.
 DR PROSITE: PS50280; SET; 1.
 DR PROSITE: PS50016; ZF_PHD_2; 1.
 DR PROSITE: PS50089; ZF_RING_2; 1.
 SQ SEQUENCE 4823 AA; 526260 MW; BD0C5FAEAD0F9CC7 CRC64;

Query Match 27.1%; Score 54; DB 13; Length 4823;
 Best Local Similarity 29.4%; Pred. No. 6.4;
 Matches 15; Conservative 2; Mismatches 20; Indels 14; Gaps 3;

OY 1 HXHXSTXXCKKFCGTACTXXC-----RXLHXG-----KXCKXHC 38
 DB 337 HHRNCTGVCKNPKDGSV-ATLQSCSVCHRLVHSGCTLPKELSEDKICILHC 386

RESULT 7
 OY 08N19 PRELIMINARY; PRT; 595 AA.
 AC 08N19;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein FLJ37933.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Tongue.
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima K., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,
 RA Isogai T.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: AK095252; BAC04510.1; -
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR007087; ZnF_C2H2.
 DR InterPro: IPR007086; ZnF_C2H2_sub.
 DR Pfam: PF01352; KRAB; 1.
 DR Pfam: PF00096; ZF_C2H2; 5.
 DR PRINTS: PR00048; ZINC_FINGER.
 DR SMART: SM00349; KRAB; 1.
 DR SMART: SM00355; ZnF_C2H2; 11.
 DR PROSITE: PS50805; KRAB; 1.
 DR PROSITE: PS50028; ZINC_FINGER_C2H2_1; 10.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 12.
 KW Hypothetical protein; Metal-binding; Nuclear protein; zinc;
 KW Zinc-finger.
 SQ SEQUENCE 595 AA; 69150 MW; E3BCG367378F7EE4 CRC64;

Query Match 26.9%; Score 53.5; DB 4; Length 595;
 Best Local Similarity 34.1%; Pred. No. 1.4;
 Matches 14; Conservative 3; Mismatches 11; Indels 13; Gaps 4;

OY 8 YXCKKFCGTACTXXC-----RXLHXG-KXCKXHCGR 40
 DB 200 YKC-KFCGKA---FHCLRLYLHRIHIGERKPCCKQCGK 235

RESULT 8
 OY 08IY18 PRELIMINARY; PRT; 595 AA.
 AC 08IY18;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein FLJ37933.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC035760; AAH35760.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 595 AA; 69105 MW; 7DCDFB98E40661 CRC64;

Query Match 26.9%; Score 53.5; DB 4; Length 595;
 Best Local Similarity 34.1%; Pred. No. 1.4;
 Matches 14; Conservative 3; Mismatches 11; Indels 13; Gaps 4;

OY 8 YXCKKFCGTACTXXC-----RXLHXG-KXCKXHCGR 40
 ID 08N19 PRELIMINARY; PRT; 595 AA.
 AC 08N19;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)

Db 200 YKC-KFCGKA---FHCLRLYLIIHRIHTGKPCCKQCKG 235

RESULT 9

08C020 PRELIMINARY; PRT; 771 AA.

AC 08C020 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Weekly similar to CDNA FLJ11726 FIS (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AK032536; BAC27915.1; -

FT NON_TER 771

SQ SEQUENCE 771 AA; 85031 MW; 824647619335E9AA CRC64;

Query Match 26.6%; Score 53; DB 11; Length 771;

Best Local Similarity 32.4%; Pred. No. 2;

Matches 11; Conservative 1; Mismatches 14; Indels 8; Gaps 1;

QY 14 CGTAXCTYXXCRL-----HXKXCKXCHCS 39

Db 180 CGVCGCSFSSCSALEKHVECHVEGKERTCCHCS 213

RESULT 10

09U0E2 PRELIMINARY; PRT; 585 AA.

AC 09U0E2;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Signal receptor protein (Fragment).

GN NOTCH.

OC Tribolium castaneum (Red flour beetle).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;

OC Tenebrionidae; Tribolium.

OX NCBI_TaxID=7070;

RN [1]

RP SEQUENCE FROM N.A.

RA Tautz D., Lardelli M., Westin J., Tamme R.;

RA "Embryonic expression of the Tribolium Notch."

RL Submitted (AFR-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL; AJ005083; CAB65469.1; -

DR HSSP; P00740; IEDM.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000742; EGF_2.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR001438; EGF_11.

DR InterPro; IPR006209; EGF_1like.

DR InterPro; IPR000800; Notch.

DR Pfam; PF00008; EGF_11.

DR Pfam; PF00066; notch_2.

DR PRINTS; PR00010; EGFBL00.

DR PRINTS; PR01452; NOTCH.

DR SMART; SM00179; EGF_CA_5.

DR SMART; SM00004; NL_2.

DR PROSITE; PS00010; ASX_HYDROXYL; 6.

DR PROSITE; PS00022; EGF_1; 12.

DR PROSITE; PS01186; EGF_2; 9.

DR PROSITE; PS01187; EGF_CA; 4.

KW EGF-like domain; Receptor.

FT NON_TER 1

FT NON_TER 585

SQ SEQUENCE 585 AA; 63337 MW; 4CF7A51D0820D048 CRC64;

Query Match 26.4%; Score 52.5; DB 5; Length 585;

Best Local Similarity 37.5%; Pred. No. 1.9;

Matches 9; Conservative 2; Mismatches 12; Indels 1; Gaps 1;

QY 13 FCGTAXC-TYXXCRLHXKXKXC 35

Db 341 FCATSPCNGVCVCTTIHAGHKCTC 364

RESULT 11

09V3J9 PRELIMINARY; PRT; 882 AA.

AC 09V3J9;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE BG:DS07295.2 protein.

GN BG:DS07295.2 OR CG15267.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OX Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,

RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,

RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cavalier S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler E., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RT Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RL "The genome sequence of Drosophila melanogaster."

[2]

RP SEQUENCE FROM N.A.

```

RC STRAIN-Berkeley;
RA MEDLINE:99403001; PubMed:10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartwell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,
RA Celniker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RA Celniker S.E., Agbayani A., Arcaena T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Champagne M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Katta K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomoclan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Smit E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AE003646; AAF53444.1;
DR EMBL: AE003412; AAF4928.1;
DR Flybase: FBgn0028862; BG:DS07295.2.
KM Hypothetical protein.
SQ SEQUENCE 882 AA; 99053 MW; B28247DA0EA4706A CRC64;

Query Match 26.4%; Score 52.5; DB 5; Length 882;
Best Local Similarity 23.8%; Pred. No. 2.7;
Matches 15; Conservative 2; Mismatches 19; Indels 27; Gaps 2;

OY 3 HXCTGYCKKFCG-----TAXCTYXC-----RXHXGKXKC 35
DB 555 HSCSTTAKSGRVTALALDLPAGSGTYNCFAGNFQQLPREERTKQLLESGIRKHC 654
OY 36 XHC 38
DB 655 NAC 657

RESULT 12
OYV525 PRELIMINARY; PRT; 603 AA.
AC OYV525;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG8083 protein (LP06581P).
DE BEST:CK00459 OR CG8083.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE:20196006; PubMed:10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Ffannkoch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Butler H., Cadieu E., Center A., Chandra I.,
RA Burtis K.C., Busan D.A., Butler J., Brokstein P., Brottier P.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```

```

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
RA Fioder A., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodok C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AE003834; AAF58997.1;
DR EMBL: AY052046; AAK93470.1;
DR Flybase: FBgn0025709; BEST:CK00459.
DR InterPro: IPR002668; NucleoSide.transp2.
DR Pfam: PF01773; NucleoSide.transp2.
DR Prodom: PD003768; NucleoSide.transp2.1.
SQ SEQUENCE 603 AA; 67179 MW; 658B6E3CA12A3A9 CRC64;

Query Match 26.1%; Score 52; DB 5; Length 603;
Best Local Similarity 35.3%; Pred. No. 2.4;
Matches 12; Conservative 1; Mismatches 17; Indels 4; Gaps 1;

OY 1 HXHXCTGYC-----KFCGTACCTYXXCXALXG 30
DB 79 HYHDTNCEYKWKNDPLGCGINECTGYGMLLLG 112

RESULT 13
OYV889 PRELIMINARY; PRT; 163 AA.
AC OYV889;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE WSV094 (WSSV150).
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE:21548311; PubMed:11689662;
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT "Complete genome sequence of the shrimp white spot bacilliform
RT virus.";
RL J. Virol. 75:11811-11820(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
RN [3]

```

[illegible]

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacled J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradlin A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL science 287:2185-2195(2000).
 [2]

RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacled J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003458; AAF46870.1; -
 DR EMBL: AY061190; AAL28738.1; -
 DR FLYBase: FBgn0034750; CG3732.
 DR InterPro: IPR001876; Znf_RanGDP.
 DR Pfam: PF00641; zif-RanBP. 1.
 DR SMART: SM00547; znf_RanBP. 2.
 DR PROSITE: PS01358; ZF_RANBP2_1; 2.
 DR PROSITE: PS0199; ZF_RANBP2_2; 2.
 SQ SEQUENCE 282 AA; 31033 MW; F916D795D315D3BA CRC64;

Query Match 25.6%; Score 51; DB 5; Length 282;
 Best Local Similarity 34.6%; Pred. No. 1.8;
 Matches 9; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 15 GTACCTYYXCXHXGXKXKXCHCSR 40
 DB 22 GDWICPDYDCRHINFARRLCNKCDCR 47

Search completed: September 3, 2003, 11:01:41
 Job time : 63.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:55:11; Search time 25 seconds

(without alignments)
153.870 Million cell updates/sec

Title: US-10-030-231-5

Perfect score: 199

Sequence: 1 HXHXCTSYXCXKFCGTACXTCYXGCRXLHXGKXCXHCSCR 40

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.5	26.4	186	2 A28401	agglutinin isolect
2	52.5	26.4	212	2 T05936	agglutinin isolect
3	52	26.1	372	1 S28296	hypothetical prote
4	51.5	23.9	72	1 T1VTOA	protease inhibit
5	50	25.1	342	2 A46529	Ig gamma chain (5.
6	49	24.6	352	2 T31544	hypothetical prote
7	49	24.6	379	2 A35669	gene CYR61 protein
8	48.5	24.4	197	2 I46413	keratin KAP5-5-s
9	48.5	24.4	606	2 S43118	finger protein - m
10	48	24.1	693	2 S49228	sodium-dependent p
11	47.5	23.9	1801	1 MMRTS	laminin beta-2 cha
12	47.5	23.6	152	2 T18975	hypothetical prote
13	47	23.6	175	2 S37649	high-sulfur kerat
14	47	23.6	261	1 WZBE01	gene 69 protein -
15	47	23.6	261	2 T42983	hypothetical prote
16	47	23.6	344	2 T46928	hypothetical prote
17	47	23.6	633	2 T17262	hypothetical prote
18	46.5	23.4	213	1 AEWT2	agglutinin isolect
19	46.5	23.4	806	2 A46271	integrin beta-7 ch
20	46.5	23.4	1895	2 T06609	disease resistance
21	46.5	23.4	3461	2 S58870	reelin precursor -
22	46.5	23.4	3712	2 S18253	laminin alpha-1 ch
23	46	23.1	164	2 T24272	hypothetical prote
24	46	23.1	188	2 T15651	hypothetical prote
25	46	23.1	236	2 T18323	hypothetical prote
26	46	23.1	367	2 S06582	finger protein (cl
27	46	23.1	425	2 JCS909	AE33 protein - fru
28	46	23.1	564	2 T12489	hypothetical prote
29	45.5	22.9	577	2 B37057	integrin beta-6 ch

30	45.5	22.9	788	2 A37057	integrin beta-6 ch
31	45.5	22.9	1955	1 AGCH	agrin precursor -
32	45	22.6	106	4 S57386	hypothetical prote
33	45	22.6	115	2 G84848	metallothionein-II
34	45	22.6	139	2 T40076	probable cyclin G1
35	45	22.6	149	2 AG0664	hypothetical prote
36	45	22.6	169	1 S18946	ultra high-sulfur
37	45	22.6	218	2 G91207	probable replicase
38	45	22.6	223	2 A65172	hypothetical prote
39	45	22.6	325	2 A86054	probable replicase
40	45	22.6	860	1 QRHULD	LDL receptor precu
41	45	22.6	940	2 T31575	hypothetical prote
42	45	22.6	1172	2 A42587	thrombospondin 2 p
43	45	22.6	2630	2 T08868	polyprotein p1 - A
44	45	22.6	4545	1 S25111	alpha-2-macroglobu
45	44.5	22.4	26	2 S55029	CAB3 protein - ant

ALIGNMENTS

RESULT 1

A28401 agglutinin isolectin 3 precursor - wheat (fragment)

N:Alternate names: agglutinin isolectin B

C:Species: Triticum aestivum (common wheat)

C>Date: 19-May-1989 #sequence_revison 19-May-1989 #text_change 24-Nov-1999

C:Accession: A28401; S10045

R:Raikhel, N.V.; Wilkins, T.A.

A:Title: Isolation and characterization of a cDNA clone encoding wheat germ agglutinin

A:Reference number: A28401

A:Accession: A28401

A:Molecule type: mRNA

A:Residues: 1-186 <RAI>

A:Cross-References: GB:J02961; MID:9170667; PIDN:AAA4257.1; PID:9170668

R:Wright, C.S.; Raikhel, N.

J. Mol. Evol. 28, 327-336, 1989

A:Title: Sequence variability in three wheat germ agglutinin isolectins: products of

A:Reference number: S07289; MID:89279931; PMID:2496868

A:Accession: S10045

A:Molecule type: protein

A:Residues: 1-171 <MRI>

C:Comment: The three isolectins associate randomly into dimers in vivo.

C:Superfamily: wheat agglutinin; hevein chitin-binding domain homology

C:Keywords: blocked amino end; dimer; duplication; glycoprotein; lectin

F:1-11/1/Product: agglutinin isolectin 3 #status experimental <MAT>

F:1-43/Domain: hevein chitin-binding domain homology <HCB1>

F:44-86/Domain: hevein chitin-binding domain homology <HCB2>

F:87-129/Domain: hevein chitin-binding domain homology <HCB3>

F:130-172/Domain: hevein chitin-binding domain homology <HCB4>

F:172-186/Domain: carboxyl-terminal propeptide #status predicted <CPRO>

F:1/Modified site: blocked amino end (Gln) (probably pyroglutamate carboxylic acid) #st

F:3-18, 12-24, 17-31, 35-40, 46-61, 55-67, 60-74, 78-83, 89-104, 98-110, 103-117, 121-126, 132-14

F:180/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.4%; Score 52.5; DB 2; Length 186;

Best Local Similarity 30.2%; Pred. No. 6.8;

Matches 13; Conservative 1; Mismatches 20; Indels 9; Gaps 2;

Qy 5 CTSY-----XCXKFCGTAYC-TYXGCRXLHXGKXCXHC 38

Db 18 CSQYGYCGMGDYGCGCGCGCGAGCMTWSKRCGSAAGCTCPNNHC 60

RESULT 2

T05936 agglutinin isolectin 1 precursor - barley

C:Species: Hordeum vulgare (barley)

C>Date: 30-Apr-1999 #sequence_revison 30-Apr-1999 #text_change 21-Jan-2000

C:Accession: T05936

R:Riemer, D.R.; Raikhel, N.V.

Plant Physiol. 91, 124-129, 1989

A:Title: Cloning and characterization of root-specific barley lectin.
A:Reference number: 215461
A:Accession: T05936
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-212 <LER>
A:Cross-references: EMBL:M29280; NID:9167070; PIDN:AAA32969.1; PID:9167071
C:Superfamily: wheat agglutinin; hevein chitin-binding domain homology
C:Keywords: lectin
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-212/Product: agglutinin isolectin 1 #status predicted <MAT>
F:27-69/Domain: hevein chitin-binding domain homology <HCB>

Query Match
Best Local Similarity 26.4%; Score 52.5; DB 2; Length 212;
Best Local Similarity 30.2%; Pred. No. 7.4;
Matches 13; Conservative 1; Mismatches 20; Indels 9; Gaps 2;

QY 5 CTST-----KCKKFCGTAXC-TXXCRXLHGKXCXKHC 38
DB 44 CSOYCGCGMGDYGCGCGCGACGYSTRKCGTQAGGKTCPPNNHC 86

RESULT 3
S28296
hypothetical protein C40H1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S28296
R:Berkas, M.
Submitted to the EMBL Data Library, December 1992
A:Reference number: S28296
A:Accession: S28296
A:Molecule type: DNA
A:Residues: 1-372 <BER>
A:Cross-references: EMBL:Z19154; NID:96650; PID:96651
C:Genetics:
A:Introns: 72/1; 109/1; 139/2; 252/3; 324/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C40H1.1; ribonucleoprotein re
F:59-139/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match
Best Local Similarity 26.1%; Score 52; DB 1; Length 372;
Best Local Similarity 43.5%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 8; Indels 4; Gaps 1;

QY 13 FCGTAXCTYXXCR---XLMXGK 31
DB 281 FCGHASCLOYCCGCGCDRMHIGK 303

RESULT 4
T1VTON
protease inhibitor (Bowman-Birk) - common vetch
C:Species: Vicia sativa subsp. nigra (common vetch)
C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 17-Feb-1995
C:Accession: A01304
R:Shimokawa, Y.; Kuroimizu, K.; Araki, T.; Ohata, J.; Abe, O.
Nat. Cult. 10, 69-73, 1983
A:Title: The complete amino acid sequence of Vicia angustifolia proteinase inhibitor.
A:Reference number: A01304
A:Accession: A01304
A:Molecule type: protein
A:Residues: 1-72 <SHI>
A:Experimental source: var. segetalis
C:Comment: One mole of inhibitor inhibits either one mole of trypsin or two moles of chy
C:Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
C:Keywords: duplication; serine proteinase inhibitor
F:9-34/Domain: Bowman-Birk inhibitor repeat homology <BB1>
F:35-60/Domain: Bowman-Birk inhibitor repeat homology <BB2>
F:8-61,9-24,12-57,14-22,31-38,35-50,40-48/Disulfide bonds: #status predicted
F:16/inhibitory site: Arg (trypsin) #status predicted
F:42/inhibitory site: Tyr (chymotrypsin) #status predicted

Query Match
25.9%; Score 51.5; DB 1; Length 72;

Best Local Similarity 36.7%; Pred. No. 5;
Matches 11; Conservative 2; Mismatches 12; Indels 5; Gaps 2;

QY 14 CGTAXCTFY---XCRXLHGKXC--XCXHC 38
DB 9 CDFCLCTRSOPPTRCVDEGRCHSACNHC 38

RESULT 5
A46529
Ig gamma chain (5.7S) - duck (fragment)
C:Species: Anas platyrhynchos (domestic duck)
C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A46529; S20760
R:Major, K.E.; Marr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.
J. Immunol. 149, 2627-2633, 1992
A:Title: Structural relationship between the two Ig of the duck, Anas platyrhynchos:
A:Reference number: A46529; MUID:93017865; PMID:1401901
A:Accession: A46529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-342 <MAG>
A:Cross-references: EMBL:X65218; NID:962444; PIDN:CAA46321.1; PID:g1334619
A:Experimental source: spleen
A>Note: sequence extracted from NCBI backbone (NCBI:116122)
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:22-104/Domain: immunoglobulin homology <IMM>

Query Match
Best Local Similarity 25.1%; Score 50; DB 2; Length 342;
Best Local Similarity 44.4%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 7 SYXCKKFCGTAXCTYXXC 24
DB 99 TTYCAKITGYANCACTGC 116

RESULT 6
T31544
hypothetical protein Y47D3A.12 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31544
R:Matthews, L.
Submitted to the EMBL Data Library, October 1999
A:Reference number: Z21043
A:Accession: T31544
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-352 <WIL>
A:Cross-references: EMBL:AL117022; PIDN:CAB57892.1; CESP:Y47D3A.12
A:Experimental source: clone Y47D3A
C:Genetics:
A:Gene: CESP:Y47D3A.12
A:Introns: 37/2; 108/2; 214/3

Query Match
Best Local Similarity 24.6%; Score 49; DB 2; Length 352;
Best Local Similarity 61.5%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 SYXCKKFCGTAXC 19
DB 3 SYVCTKRLCPATAC 15

RESULT 7
A35669
gene CYP61 protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 05-Nov-1999
C:Accession: A35669; S16446
R:O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.

MOL. Cell. Biol. 10, 3569-3577, 1990
 A:Title: Expression of cyr61, a growth factor-inducible immediate-early gene.
 A:Reference number: A35669; MUID:90287146; PMID:3355916
 A:Accession: A35669
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-379 <OAB>
 A:Cross-references: GB:M32490; NID:9192909; PIDN:AAA37512.1; PID:9309206
 A:Note: the authors translated the codon GAT for residue 337 as Gln
 R:Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.
 Nucleic Acids Res. 19, 3261-3267, 1991
 A:Title: Promoter function and structure of the growth factor-inducible immediate early
 A:Reference number: 148319; MUID:91288203; PMID:2062642
 A:Accession: 148319
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-379 <RES>
 A:Cross-references: EMBL:X56790; NID:950632; PIDN:CAA40109.1; PID:950633
 A:Note: the authors did not translate the codon for residue 108
 A:Note: the authors translated the codon GAT for residue 337 as Gln
 C:Genetics:
 A:Gene: CYR61
 A:Introns: 21/3; 93/1; 208/1; 279/3
 C:Superfamily: von Willebrand factor type C repeat homology
 F:99-166/Domain: von Willebrand factor type C repeat homology <VMC>
 Query Match 24.6%; Score 49; DB 2; Length 379;
 Best Local Similarity 28.0%; Pred. No. 29;
 Matches 14; Conservative 4; Mismatches 22; Indels 10; Gaps 3;

QY 1 HXHC---TSYX-CKKFCGTAXCTYXX---CXRLHXGKXCXCHCSR 40
 DB 223 HGQKIVCTTSMSCSGSGTISRTVINDNECHLVETRICVPRCQ 272
 16413
 keratin KAP5.5 - sheep (fragment)
 C:Species: Ovis orientalis aries; Ovis ammon aries (domestic sheep)
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 24-Sep-1999
 C:Accession: 146413; S34216
 R:Jenkins, B.J.; Powell, B.C.
 J. Invest. Dermatol. 103, 310-317, 1994
 A:Title: Differential expression of genes encoding a cysteine-rich keratin family in the
 A:Reference number: 146412; MUID:94358466; PMID:7521375
 A:Accession: 146413
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-197 <JEN>
 A:Cross-references: EMBL:X73435; NID:9313721; PIDN:CAA51830.1; PID:9313722
 C:Genetics:
 A:Gene: KRTAP5.5
 C:Superfamily: ultra-high-sulfur keratin
 Query Match 24.4%; Score 48.5; DB 2; Length 197;
 Best Local Similarity 29.0%; Pred. No. 23;
 Matches 9; Conservative 3; Mismatches 14; Indels 5; Gaps 1;

QY 10 CKKFCGTAXCTYXXCRXLHXGKXCXCHCSR 40
 DB 37 CGSGCGSCCVVC-----VPACSCSCSK 62
 54318
 finger protein - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
 C:Accession: 54318
 R:Brady, J.P.; Piatigorsky, J.
 submitted to the EMBL Data Library, March 1994
 A:Description: A novel cDNA clone isolated from the mouse eye lens has zinc fingers and
 A:Reference number: 54318

A:Accession: S43118
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-606 <BRA>
 A:Cross-references: GB:U07861; EMBL:Z31370; NID:9466390; PIDN:AAB39204.1; PID:9466391
 C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
 Query Match 24.4%; Score 48.5; DB 2; Length 606;
 Best Local Similarity 35.1%; Pred. No. 46;
 Matches 13; Conservative 3; Mismatches 16; Indels 5; Gaps 3;

QY 8 YXCKFCGTAXCTYXXCRX---LHXG-KXCKXCHCSR 40
 DB 492 YAC-KHCGKAFPTSSARNSHERIHTEKPYTCKHCTK 527
 549228
 sodium-dependent phosphate transporter - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 05-Nov-1999
 C:Accession: S68972; S49228
 R:Helms, C.; Murer, H.; McGivan, J.
 Eur. J. Biochem. 228, 927-930, 1995
 A:Title: Cloning, sequence analysis and expression of the cDNA encoding a sodium-depe
 A:Reference number: S68972; MUID:9525303; PMID:7737195
 A:Accession: S68972
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-693 <HE2>
 A:Cross-references: EMBL:X81699; NID:9547483; PIDN:CAA57345.1; PID:9547484
 Query Match 24.1%; Score 48; DB 2; Length 693;
 Best Local Similarity 28.6%; Pred. No. 57;
 Matches 10; Conservative 1; Mismatches 14; Indels 10; Gaps 1;

QY 5 CTSYXCKKFCGTAXCTYXXCRXLHXGKXCXCHCS 39
 DB 620 CCRVCCRLCCGCGCS-----KCCRCKCS 644
 620
 laminin beta-2 chain precursor - rat
 M:MRNS
 N:Alternate names: laminin chain B3; S-laminin
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-Dec-1999
 C:Accession: S03539
 R:Hunter, D.D.; Shah, V.; Merlie, J.P.; Sanes, J.R.
 Nature 338, 229-234, 1989
 A:Title: A laminin-like adhesive protein concentrated in the synaptic cleft of the ne
 A:Reference number: S03539; MUID:89159410; PMID:2922051
 A:Accession: S03539
 A:Molecule type: mRNA
 A:Residues: 1-1801 <HUN>
 A:Cross-references: EMBL:X16563; NID:957250; PIDN:CAA34561.1; PID:957251
 C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type lamin
 C:Function: interact with cells and with other basement membrane proteins to promo
 C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
 C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellu
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:36-1801/Product: laminin beta-2 chain #status predicted <MAT>
 F:36-289/Domain: VI <DOM6>
 F:286-555/Domain: V <DOM5>
 F:286-347/Domain: laminin-type EGF-like homology <LE01>
 F:350-410/Domain: laminin-type EGF-like homology <LE02>
 F:413-470/Domain: laminin-type EGF-like homology <LE03>
 F:473-522/Domain: laminin-type EGF-like homology <LE04>
 F:525-555/Domain: laminin-type EGF-like homology #status atypical <LE05>
 F:556-784/Domain: IV <DOM4>
 F:786-833/Domain: laminin-type EGF-like homology <LE06>
 F:788-1196/Domain: III <DOM3>

F:834-877/Domain: laminin-type EGF-like homology <LE07>
 F:880-927/Domain: laminin-type EGF-like homology <LE08>
 F:930-986/Domain: laminin-type EGF-like homology <LE09>
 F:989-1038/Domain: laminin-type EGF-like homology <LE10>
 F:1041-1095/Domain: laminin-type EGF-like homology <LE11>
 F:1098-1143/Domain: laminin-type EGF-like homology <LE12>
 F:1146-1190/Domain: laminin-type EGF-like homology <LE13>
 F:1197-1412/Domain: II <DOM2>
 F:1197-1412/Region: heptad repeats
 F:1413-1445/Domain: alpha <ALP>
 F:1446-1801/Region: heptad repeats
 F:1446-1801/Domain: I <DOM1>
 F:445-50/Disulfide bonds: #status predicted
 F:251,371,1088,1252,1311,1351,1502/Binding site: carbohydrate (Asn) (covalent) #status F
 F:1193,1196,1800/Disulfide bonds: interchain #status predicted

Query Match 23.9%; Score 47.5; DB 1; Length 1801;
 Best Local Similarity 24.6%; Pred. No. 1.2e+02;
 Matches 14; Conservative 3; Mismatches 13; Indels 27; Gaps 3;

OY 10 CXXKCGTAXC-----TXXYCKXHXG-----KXKC-----XHC 39
 DB 1110 CNEFTGQCHAGFGGRTCSQELHWDPELQCRACDCDPRGIDKPCQCHRSTGHCS 1166

RESULT 12

hypothetical protein C06A1.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18975

R:McMurray, A.
 submitted to the EMBL Data Library, June 1995

A:Reference number: Z19054

A:Accession: T18975

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-152 <M12>

A:Cross-references: EMBL:Z49886; PIDD:CAA90055.1; GSPDB:GN00020; CESP:C06A1.6

A:Experimental source: clone C06A1

C:Genetics:

A:Gene: CESP:C06A1.6

A:Map position: 2

A:Introns: 22/3

Query Match 23.6%; Score 47; DB 2; Length 152;
 Best Local Similarity 35.3%; Pred. No. 30;

Matches 12; Conservative 0; Mismatches 20; Indels 2; Gaps 1;

OY 5 CTSYXCKRFGCTAXCTYXCRXLHXGKXCXHC 38
 DB 86 CCCTCCRTCCCTRCCT--CCRPCCGCGCGCGCC 117

RESULT 13

S37649

high-sulfur keratin - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999

C:Accession: S37649

R:Rhumabaeva, B.D.; Genting, L.V.; Gazaryan, K.G.

Mol. Biol. 26, 550-555, 1992

A:Title: Cloning and structural characterization of human hair sulfur-rich keratin genes

A:Reference number: S37649

A:Accession: S37649

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-175 <RHU>

A:Cross-references: EMBL:X63338; NID:g311881; PIDD:CAA44938.1; PID:g311882

C:Superfamily: keratin high-sulfur matrix protein IIA

Query Match 23.6%; Score 47; DB 2; Length 175;
 Best Local Similarity 25.0%; Pred. No. 33;

Matches 12; Conservative 2; Mismatches 20; Indels 14; Gaps 2;

OY 5 CTSYXCKRFGCTAXCTYX-----CRXLHXGKC-----XCHC 38
 DB 128 CLRPCCVCSCHTTPCCOLHHAASCRPSYCSGSCCPVCCYCSCHC 175

RESULT 14

gene 69 protein - salmeline herpesvirus 1 (strain 11)

C:Species: salmeline herpesvirus 1

A:Note: host Salmi scutours (common squirrel monkey)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999

C:Accession: D36813

R:Albrecht, J.

submitted to the EMBL Data Library, January 1992

A:Description: Primary structure of the herpesvirus salm1 genome.

A:Reference number: A36806

A:Accession: D36813

A:Molecule type: DNA

A:Residues: 1-261 <ALB>

A:Cross-references: GB:X64346; NID:g60320; PIDD:CAA45692.1; PID:g60390

R:Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.;

J. Virol. 66, 5047-5058, 1992

A:Title: Primary structure of the herpesvirus salm1 genome.

A:Reference number: A37309; MUID:92333688; PMID:1321287

A:Contents: annotation; possible protein-coding frames

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 69

C:Superfamily: varicella-zoster virus gene 27 protein

Query Match 23.6%; Score 47; DB 1; Length 261;
 Best Local Similarity 40.0%; Pred. No. 42;

Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 19 CTYXCRXLHXGKXC 33
 DB 175 CLTFKQTLHIGESC 189

RESULT 15

T42983

hypothetical protein 69 - ateline herpesvirus 3 (strain 73)

C:Species: ateline herpesvirus 3

A:Variety: strain 73

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T42983

R:Albrecht, J.C.; Fleckenstein, B.

submitted to the EMBL Data Library, August 1998

A:Description: Primary structure of the herpesvirus ateles genome.

A:Reference number: Z22274

A:Accession: T42983

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-261 <ALB>

A:Cross-references: EMBL:AF083424; PIDD:AMC95594.1

A:Experimental source: strain 73

C:Superfamily: varicella-zoster virus gene 27 protein

Query Match 23.6%; Score 47; DB 2; Length 261;
 Best Local Similarity 40.0%; Pred. No. 42;

Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 19 CTYXCRXLHXGKXC 33
 DB 175 CLTFKQTLHIGESC 189

Search completed: September 3, 2003, 11:02:36
 Job time : 26.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:59:06 : Search time 19.5 Seconds
(without alignments)
281.649 Million cell updates/sec

Title: US-10-030-231-5

Perfect score: 199

Sequence: 1 HXHXCTSYXCKKFCGTAXCTYXCRLHXKXCXCHCSR 40

Scoring table:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCITUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	33.2	2095	US-10-184-644-161	Sequence 161, App
2	66	33.2	2095	US-10-184-634-161	Sequence 161, App
3	65	32.7	2089	US-10-140-472-497	Sequence 497, App
4	65	32.7	2089	US-10-141-761-497	Sequence 497, App
5	65	32.7	2089	US-10-142-885-497	Sequence 497, App
6	65	32.7	2089	US-10-123-155-497	Sequence 497, App
7	65	32.7	2089	US-10-146-731-497	Sequence 497, App
8	63	31.7	2732	US-09-086-436-90	Sequence 473, App
9	63	31.7	2870	US-10-184-644-473	Sequence 473, App
10	63	31.7	2870	US-10-184-634-473	Sequence 473, App
11	63	31.7	3244	US-10-184-644-571	Sequence 571, App
12	63	31.7	3244	US-10-184-634-571	Sequence 571, App
13	63	31.7	3608	US-10-184-644-433	Sequence 433, App
14	63	31.7	3608	US-10-184-634-433	Sequence 433, App
15	62	31.2	1497	US-09-060-854B-2	Sequence 2, Appl

16	62	31.2	3907	US-10-029-217A-24	Sequence 24, Appl
17	61.5	30.9	3479	US-10-140-472-123	Sequence 123, App
18	61.5	30.9	3479	US-10-141-761-123	Sequence 123, App
19	61.5	30.9	3479	US-10-142-885-123	Sequence 123, App
20	61.5	30.9	3479	US-10-123-155-123	Sequence 123, App
21	61.5	30.9	3479	US-10-146-731-123	Sequence 123, App
22	61.5	30.9	4679	US-09-804-898-2	Sequence 2, Appl
23	61	30.7	636	US-10-140-472-507	Sequence 507, App
24	61	30.7	636	US-10-141-761-507	Sequence 507, App
25	61	30.7	636	US-10-142-885-507	Sequence 507, App
26	61	30.7	636	US-10-123-155-507	Sequence 507, App
27	61	30.7	1505	US-10-146-731-507	Sequence 507, App
28	61	30.7	1505	US-10-184-644-377	Sequence 377, App
29	61	30.7	1505	US-10-184-634-377	Sequence 377, App
30	61	30.7	2168	US-10-184-644-43	Sequence 43, Appl
31	61	30.7	2168	US-10-184-634-43	Sequence 43, Appl
32	61	30.7	2747	US-10-184-644-101	Sequence 101, App
33	61	30.7	2747	US-10-184-634-101	Sequence 101, App
34	60.5	30.4	1660	US-10-184-644-147	Sequence 147, App
35	60.5	30.4	1660	US-10-184-634-147	Sequence 147, App
36	60	30.2	1174	US-10-184-644-353	Sequence 353, App
37	60	30.2	1174	US-10-184-634-353	Sequence 353, App
38	60	30.2	1319	US-10-140-472-241	Sequence 241, App
39	60	30.2	1319	US-10-141-761-241	Sequence 241, App
40	60	30.2	1319	US-10-142-885-241	Sequence 241, App
41	60	30.2	1319	US-10-123-155-241	Sequence 241, App
42	60	30.2	1319	US-10-146-731-241	Sequence 241, App
43	60	30.2	2290	US-10-140-472-267	Sequence 267, App
44	60	30.2	2290	US-10-141-761-267	Sequence 267, App
45	60	30.2	2290	US-10-142-885-267	Sequence 267, App

ALIGNMENTS

RESULT 1
US-10-184-644-161
: Sequence 161, Application US/10184644
: Publication No. US20030044930A1
GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 161
LENGTH: 2095
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-161

Query Match 33.2% Score 66; DB 15; Length 2095;
Best Local Similarity 34.3% Pred. No. 2.2;
Matches 12; Conservative 2; Mismatches 21; Indels 0; Gaps 0;

QY 5 CTSYXCKKFCGTAXCTYXCRLHXKXCXCHCS 39
DB 995 CTACACAGAGATATCTTCAAGGTGTCCTCA 1029

```
RESULT 2
US-10-184-634-161
; Sequence 161, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 161
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-161

Query Match          33.2%; Score 66; DB 15; Length 2095;
Best Local Similarity 34.3%; Pred. No. 2.2;
Matches 12; Conservative 2; Mismatches 21; Indels 0; Gaps 0;

QY 5 CTSTYXCKRFGCTAYACTYXXCRXLHGXKXCXHC 39
DB 995 CTTACCAGAGATCTTCAAGTGCTCCCTCCA 1029

RESULT 3
US-10-140-472-497
; Sequence 497, Application US/10140472
; Publication No. US20030138888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 497
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-472-497
```

```
Query Match          32.7%; Score 65; DB 12; Length 2089;
Best Local Similarity 35.5%; Pred. No. 2.9;
Matches 11; Conservative 1; Mismatches 19; Indels 0; Gaps 0;

QY 5 CTSTYXCKRFGCTAYACTYXXCRXLHGXKXCX 35
DB 1165 CTTCTCACCCCTTGTGCTGACCTTAGTGAACCTC 1195

RESULT 4
US-10-141-761-497
; Sequence 497, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; PRIOR APPLICATION REMOVED - See File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 497
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-141-761-497

Query Match          32.7%; Score 65; DB 12; Length 2089;
Best Local Similarity 35.5%; Pred. No. 2.9;
Matches 11; Conservative 1; Mismatches 19; Indels 0; Gaps 0;

QY 5 CTSTYXCKRFGCTAYACTYXXCRXLHGXKXCX 35
DB 1165 CTTCTCACCCCTTGTGCTGACCTTAGTGAACCTC 1195

RESULT 5
US-10-142-885-497
; Sequence 497, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/142,885
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 497
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-885-497
```

APPLICANT: Wood,William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C248
CURRENT FILING DATE: 2002-05-10
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 497
LENGTH: 2089
TYPE: DNA
ORGANISM: Homo Sapien
US-10-142-885-497

Query Match 32.7%; Score 65; DB 12; Length 2089;
Best Local Similarity 35.5%; Pred. No. 2.9;
Matches 11; Conservative 1; Mismatches 19; Indels 0; Gaps 0;

Qy 5 CTGYXCKRFGCTAXCTYXXCRXLHXGKXCX 35
Db 1165 CTCTCACCCCTTGTCTGACCTTAGTGAACTC 1195

RESULT 6
US-10-123-155-497

Sequence 497, Application US/10123155
Publication No. US20030068794A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 497
LENGTH: 2089
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-497

Query Match 32.7%; Score 65; DB 15; Length 2089;
Best Local Similarity 35.5%; Pred. No. 2.9;
Matches 11; Conservative 1; Mismatches 19; Indels 0; Gaps 0;

Qy 5 CTGYXCKRFGCTAXCTYXXCRXLHXGKXCX 35
Db 1165 CTCTCACCCCTTGTCTGACCTTAGTGAACTC 1195

RESULT 7
US-10-146-731-497
Sequence 497, Application US/10146731
Publication No. US20030129692A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C323
CURRENT APPLICATION NUMBER: US/10/146,731
CURRENT FILING DATE: 2002-05-15
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 497
LENGTH: 2089
TYPE: DNA
ORGANISM: Homo Sapien
US-10-146-731-497

Query Match 32.7%; Score 65; DB 16; Length 2089;
Best Local Similarity 35.5%; Pred. No. 2.9;
Matches 11; Conservative 1; Mismatches 19; Indels 0; Gaps 0;

Qy 5 CTGYXCKRFGCTAXCTYXXCRXLHXGKXCX 35
Db 1165 CTCTCACCCCTTGTCTGACCTTAGTGAACTC 1195

RESULT 8
US-09-086-436-30

Sequence 30, Application US/09086436
Publication No. US20030118988A1
GENERAL INFORMATION:
APPLICANT: Kandell, Eric R.
APPLICANT: Santoro, Bina
APPLICANT: Bartsch, Dusan
APPLICANT: Siegelbaum, Steven
APPLICANT: Tibbs, Gareth
TITLE OF INVENTION: Brain or Heart Cyclic Nucleotide Gated Ion Channel and
FILE REFERENCE: 0575/54806-A
CURRENT APPLICATION NUMBER: US/09/086,436
CURRENT FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 30
LENGTH: 2732
TYPE: PRT
ORGANISM: Murine
US-09-086-436-30

Query Match 31.7%; Score 63; DB 11; Length 2732;
Best Local Similarity 31.4%; Pred. No. 6.5;
Matches 11; Conservative 2; Mismatches 22; Indels 0; Gaps 0;

Qy 5 CTGYXCKRFGCTAXCTYXXCRXLHXGKXCXCHCS 39
Db 2425 CCTCGACGCTTCTCTGCCCCATGAGGTCTCCACT 2459

RESULT 9

US-10-184-644-473
; Sequence 473, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 473
; LENGTH: 2870
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-473

Query Match
Best Local Similarity 34.3%; Score 63; DB 15; Length 2870;
Matches 12; Conservative 1; Mismatches 18; Indels 4; Gaps 1;

QY 5 CTSTYXCKKFCGTACTYXXCRXLHXGKXCXCHCS 39
DB 831 CTGAGCTGTCGTGTTTGAC---AGACCTCAGCT 861

RESULT 10
US-10-184-634-473
; Sequence 473, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 473
; LENGTH: 2870
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-473

Query Match
Best Local Similarity 31.7%; Score 63; DB 15; Length 2870;
Matches 12; Conservative 1; Mismatches 18; Indels 4; Gaps 1;

QY 5 CTSTYXCKKFCGTACTYXXCRXLHXGKXCXCHCS 39
DB 831 CTGAGCTGTCGTGTTTGAC---AGACCTCAGCT 861

US-10-184-644-571
; Sequence 571, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 571
; LENGTH: 3244
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-571

Query Match
Best Local Similarity 31.7%; Score 63; DB 15; Length 3244;
Matches 11; Conservative 2; Mismatches 21; Indels 0; Gaps 0;

QY 6 TSYXCKKFCGTACTYXXCRXLHXGKXCXCHCS 39
DB 3002 TTTCACACCTTTCTTACTTCATGTCCTCCATCA 3035

RESULT 12
US-10-184-634-571
; Sequence 571, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 571
; LENGTH: 3244
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-571

Query Match
Best Local Similarity 31.7%; Score 63; DB 15; Length 3244;
Matches 11; Conservative 2; Mismatches 21; Indels 0; Gaps 0;

OY 6 TSYXCKRFGTAXCTYXXCRXLHGKXCXCHCS 39
DB 3002 TTGCACACCTTCTTACTCATGTCCCATCA 3035

RESULT 13

US-10-184-644-433
; Sequence 433, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 433
; LENGTH: 3608
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-433

Query Match 31.7%; Score 63; DB 15; Length 3608;
Best Local Similarity 35.3%; Pred. No. 8.2;
Matches 12; Conservative 1; Mismatches 17; Indels 4; Gaps 1;

OY 6 TSYXCKRFGTAXCTYXXCRXLHGKXCXCHCS 39
DB 1006 TCAGCATCCGTACACGAC---TGCACACACGA 1035

RESULT 14

US-10-184-634-433
; Sequence 433, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 433
; LENGTH: 3608
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-433

Query Match 31.7%; Score 63; DB 15; Length 3608;

Best Local Similarity 35.3%; Pred. No. 8.2;
Matches 12; Conservative 1; Mismatches 17; Indels 4; Gaps 1;

OY 6 TSYXCKRFGTAXCTYXXCRXLHGKXCXCHCS 39
DB 1006 TCAGCATCCGTACACGAC---TGCACACACGA 1035

RESULT 15

US-09-060-854B-2
; Sequence 2, Application US/09060854B
; Patent No. US20020081703A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David Aaron
; TITLE OF INVENTION: Human Protease and Use of Such Protease for Pharmaceutical
; TITLE OF INVENTION: Applications and for Reducing the Allergenicity of No. US2002
; FILE REFERENCE: GC532
; CURRENT APPLICATION NUMBER: US/09/060,854B
; CURRENT FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1497
; TYPE: PRT
; ORGANISM: B. amyloliquefaciens
US-09-060-854B-2

Query Match 31.2%; Score 62; DB 9; Length 1497;
Best Local Similarity 41.4%; Pred. No. 5.2;
Matches 12; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

OY 10 CKRFGTAXCTYXXCRXLHGKXCXCHC 38
DB 1018 CTGGCGTATCTATCCAA--AGCAGCGCTTC 1044

Search completed: September 3, 2003, 11:04:09
Job time : 19.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:57:11 : Search time 20.5 Seconds
(without alignments)
82.558 Million cell updates/sec

Title: US-10-030-231-5
Perfect score: 199
Sequence: 1 HXHKCTSYXCKFCGTAXCTYXCRXLHXGKXCXCHCSR 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/CTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	28.6	109	2	US-08-527-044-2
2	57	28.6	109	3	US-09-013-780-2
3	55	27.6	430	3	US-08-997-897-2
4	55	27.6	430	3	US-08-997-897-2
5	55	27.6	2088	4	US-09-548-372D-13
6	55	27.6	2088	4	US-09-548-372D-13
7	55	27.6	2088	4	US-09-548-372D-13
8	54.5	27.4	1917	4	US-09-627-650B-5
9	54.5	27.4	1917	4	US-09-627-650B-5
10	54.5	27.4	1917	4	US-09-627-650B-5
11	53.5	26.9	2508	4	US-09-436-063C-5
12	53.5	26.9	2508	4	US-09-436-063C-5
13	53.5	26.9	2508	4	US-09-436-063C-5
14	53.5	26.9	2508	4	US-09-436-063C-5
15	53.5	26.9	2508	4	US-09-436-063C-5
16	53.5	26.9	2508	4	US-09-436-063C-5
17	53.5	26.9	2508	4	US-09-436-063C-5
18	51.5	25.9	908	4	US-08-900-230-3
19	51.5	25.9	908	4	US-08-900-230-3
20	50	25.1	143	3	US-08-738-884-1
21	50	25.1	143	3	US-08-990-823-112
22	50	25.1	143	3	US-09-477-135A-112
23	50	25.1	1652	4	US-09-627-650B-1
24	49.5	24.9	1128	4	US-09-436-063C-1
25	49.5	24.9	1128	4	US-09-627-650B-11
26	49.5	24.9	1128	4	US-09-436-063C-11
27	49	24.6	379	1	US-08-468-847B-11
27	49	24.6	379	1	US-09-142-569-2

28	48	24.1	414	4	US-09-252-991A-31151	Sequence 31151, A
29	48	24.1	1400	4	US-08-630-915A-37	Sequence 37, Appl
30	47.5	23.9	1345	2	US-08-977-767-3	Sequence 3, Appl
31	47	23.6	41	4	US-09-194-139-4	Sequence 4, Appl
32	47	23.6	149	1	US-08-319-590-25	Sequence 25, Appl
33	47	23.6	149	1	US-08-487-001A-25	Sequence 25, Appl
34	47	23.6	149	2	US-08-630-822A-25	Sequence 25, Appl
35	47	23.6	149	2	US-08-711-905-25	Sequence 25, Appl
36	47	23.6	149	2	US-09-005-069-25	Sequence 25, Appl
37	47	23.6	155	1	US-08-487-001A-35	Sequence 35, Appl
38	47	23.6	155	2	US-08-630-822A-35	Sequence 35, Appl
39	47	23.6	155	2	US-09-005-069-35	Sequence 35, Appl
40	47	23.6	158	1	US-08-319-590-26	Sequence 26, Appl
41	47	23.6	158	1	US-08-487-001A-26	Sequence 26, Appl
42	47	23.6	158	2	US-08-630-822A-26	Sequence 26, Appl
43	47	23.6	158	2	US-08-711-905-26	Sequence 26, Appl
44	47	23.6	158	2	US-09-005-069-26	Sequence 26, Appl
45	46.5	23.4	50	4	US-08-900-230-58	Sequence 58, Appl

ALIGNMENTS

```
RESULT 1
US-08-527-044-2
: Sequence 2, Application US/08527044
: Patent No. 5885568
:
: GENERAL INFORMATION:
: APPLICANT: BOWSTEAD, Janene Marilyn J.M.
: APPLICANT: TOMLEY, Fiona Margaret F.
: APPLICANT: DUNN, Patrick Paul James P.
: APPLICANT: VERMEULEN, Arnoldus Nicolaas A.N.
: TITLE OF INVENTION: Coccidiosis poultry vaccine
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESSES:
: ADDRESSER: Akzo NO. 5885568el N.V
: STREET: 1300 PICCARD DRIVE #206
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850-4373
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/527,044
: FILING DATE: 12-SEP-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: BP 94202667.6
: FILING DATE: 16-SEP-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: GORMLEY, MARY E.
: REGISTRATION NUMBER: 34,409
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 109 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-527-044-2
:
: Query Match 28.6%; Score 57; DB 2; Length 109;
: Best Local Similarity 30.6%; Pred. No. 2.1;
: Matches 11; Conservative 4; Mismatches 17; Indels 4; Gaps 1;
:
: QY 5 CTSYXCKFCGTAXCTYXCRXLHXGKXCXCHCSR 40
: 1::| | | | | | | | | | | | | | | |
: Db 40 CSTYXCKFCGTAXCTYXCRXLHXGKXCXCHCSR 40
:
: RESULT 2
```

US-09-013-780-2
; Sequence 2, Application US/09013780
; Patent No. 6001363
; GENERAL INFORMATION:
; APPLICANT: BOMSTEAD, Janene Marilyn J.M.
; APPLICANT: TOMLEY, Fiona Margaret F.
; APPLICANT: DUNN, Patrick Paul James P.
; APPLICANT: VERMEULEN, Arnoldus Nicolaas A.N.
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 6001363el N.V
; STREET: 1300 PICCARD DRIVE #206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850-4373
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,780
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/527,044
; FILING DATE: 12-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GORMLEY, MARY E.
; REGISTRATION NUMBER: 34,409
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-013-780-2
Query Match 28.6%; Score 57; DB 3; Length 109;
Best Local Similarity 30.6%; Pred. No. 2.1;
Matches 11; Conservative 4; Mismatches 17; Indels 4; Gaps 1;
QY 5 CTSYXCKMECGTACTYXXCRXLHXGKXCXCHCSR 40
DB 40 CSYICSTFCCKCCCKCCSCFCNRSNRC---CSR 71
RESULT 3
US-08-997-897-2
; Sequence 2, Application US/08997897C
; Patent No. 6114514
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, RANJANA
; APPLICANT: KUMAR, DEEPAK
; APPLICANT: SRIVASTAVA, BRAHM SHANKER
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS SPECIFIC DNA FRAGMENT
; FILE REFERENCE: u011469-7
; CURRENT APPLICATION NUMBER: US/08/997,897C
; CURRENT FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (4)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (6)
; FEATURE:

NAME/KEY: UNSURE
LOCATION: (20)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (29)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (54)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (64)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (69)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (89)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (99)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (114)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (119)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (129)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (139)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (169)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (182)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (185)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (219)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (259)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (269)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (291)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (323)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (339)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (349)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (356)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (366)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (410)
FEATURE:
NAME/KEY: UNSURE

LOCATION: (421)
US-08-997-897-2
Query Match 27.6%; Score 55; DB 3; Length 430;
Best Local Similarity 35.1%; Pred. No. 11;
Matches 13; Conservative 2; Mismatches 20; Indels 2; Gaps 2;
Qy 5 CTSYXCKF-CGTAXCTYXXCRXLHGKXXCXCHCSR 40
Db 146 CPSLRCRFRPCRXRCQRYWCPN-XTGRCCRCPPSSR 181
RESULT 4
US-09-156-836B-2
Sequence 2, Application US/09156836B
Patent No. 6242585
GENERAL INFORMATION:
APPLICANT: Srivastava, Ranjana
APPLICANT: Kumar, Deepak
APPLICANT: Srivastava, Brahm Shanker
TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS SPECIFIC DNA FRAGMENT
FILE REFERENCE: U 011876-4
CURRENT APPLICATION NUMBER: US/09/156,836B
CURRENT FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 08/997,897
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 430
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
FEATURE:
NAME/KEY: UNSURE
LOCATION: (4)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (6)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (20)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (29)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (54)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (64)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (69)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (89)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (99)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (114)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (119)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (129)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (159)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE

LOCATION: (169)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (182)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (185)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (219)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (259)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (269)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (291)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (323)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (339)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (349)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (356)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (366)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (410)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (421)
OTHER INFORMATION: amino acid has not been identified
US-09-156-836B-2
Query Match 27.6%; Score 55; DB 3; Length 430;
Best Local Similarity 35.1%; Pred. No. 11;
Matches 13; Conservative 2; Mismatches 20; Indels 2; Gaps 2;
Qy 5 CTSYXCKF-CGTAXCTYXXCRXLHGKXXCXCHCSR 40
Db 146 CPSLRCRFRPCRXRCQRYWCPN-XTGRCCRCPPSSR 181
RESULT 5
US-09-548-372D-13
Sequence 13, Application US/09548372D
Patent No. 6420534
GENERAL INFORMATION:
APPLICANT: GURNEY ET AL.
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE. APP SUBSTRATES THEREFOR AND USE
FILE REFERENCE: 29915/62801
CURRENT APPLICATION NUMBER: US/09/548,372D
CURRENT FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: US 60/155,493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 09/404,133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: PCT/US99/20881
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 60/101,594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: PatentIn version 3.1

SEQ ID NO 13
LENGTH: 2088
TYPE: PRT
ORGANISM: Homo sapiens
US-09-548-372D-13

Query Match 27.6%; Score 55; DB 4; Length 2088;
Best Local Similarity 36.1%; Pred. No. 43;
Matches 13; Conservative 1; Mismatches 16; Indels 6; Gaps 2;

OY 5 CTSYXCKFC-GTAXCTYXCRXLHXGKXCKXCHCS 39
DB 1211 CTCGACGCTGTCTCTCTC-----GGCCTCGTCA 1241

RESULT 6
US-09-548-367D-13
Sequence 13, Application US/09548367D
Patent No. 6440698
GENERAL INFORMATION:
APPLICANT: GURNEY ET AL.
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
FILE REFERENCE: 29915/6280H
CURRENT APPLICATION NUMBER: US/09/548,367D
CURRENT FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: US 60/155,493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 09/404,133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: PCT/US99/20881
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 60/101,594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: Patentin version 3.1
SEQ ID NO 13
LENGTH: 2088
TYPE: PRT
ORGANISM: Homo sapiens
US-09-548-367D-13

Query Match 27.6%; Score 55; DB 4; Length 2088;
Best Local Similarity 36.1%; Pred. No. 43;
Matches 13; Conservative 1; Mismatches 16; Indels 6; Gaps 2;

OY 5 CTSYXCKFC-GTAXCTYXCRXLHXGKXCKXCHCS 39
DB 1211 CTCGACGCTGTCTCTCTC-----GGCCTCGTCA 1241

RESULT 7
US-09-551-853D-13
Sequence 13, Application US/09551853D
Patent No. 6500667
GENERAL INFORMATION:
APPLICANT: GURNEY ET AL.
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
FILE REFERENCE: 29915/6280L
CURRENT APPLICATION NUMBER: US/09/551,853D
CURRENT FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 60/155,493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 09/404,133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: PCT/US99/20881
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 60/101,594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: Patentin version 3.1
SEQ ID NO 13

LENGTH: 2088
TYPE: PRT
ORGANISM: Homo sapiens
US-09-551-853D-13

Query Match 27.6%; Score 55; DB 4; Length 2088;
Best Local Similarity 36.1%; Pred. No. 43;
Matches 13; Conservative 1; Mismatches 16; Indels 6; Gaps 2;

OY 5 CTSYXCKFC-GTAXCTYXCRXLHXGKXCKXCHCS 39
DB 1211 CTCGACGCTGTCTCTCTC-----GGCCTCGTCA 1241

RESULT 8
US-09-627-650B-5
Sequence 5, Application US/09627650B
Patent No. 6406872
GENERAL INFORMATION:
APPLICANT: Bamber, Bruce
APPLICANT: Jorgensen, Erik
TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
FILE REFERENCE: 21101.000903
CURRENT APPLICATION NUMBER: US/09/627,650B
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 09/436,063
PRIOR FILING DATE: 1999-11-08
PRIOR APPLICATION NUMBER: 60/107,727
PRIOR FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 1917
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-627-650B-5

Query Match 27.4%; Score 54.5; DB 4; Length 1917;
Best Local Similarity 34.3%; Pred. No. 46;
Matches 12; Conservative 2; Mismatches 14; Indels 7; Gaps 2;

OY 5 CTSYXCKFCGTAXCTYXCRXLHXGKXCKXCHCS 39
DB 1016 CTATTCAAGCGTACCA--TC-----GGCTTCTACA 1043

RESULT 9
US-09-436-063C-5
Sequence 5, Application US/09436063C
Patent No. 6407210
GENERAL INFORMATION:
APPLICANT: Bamber, Bruce
APPLICANT: Jorgensen, Erik
TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
FILE REFERENCE: P-1095corrected
CURRENT APPLICATION NUMBER: US/09/436,063C
CURRENT FILING DATE: 1999-11-08
PRIOR APPLICATION NUMBER: 60/107727
PRIOR FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 1917
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-436-063C-5

Query Match 27.4%; Score 54.5; DB 4; Length 1917;
Best Local Similarity 34.3%; Pred. No. 46;
Matches 12; Conservative 2; Mismatches 14; Indels 7; Gaps 2;

```

Oy      5      CTSYCKKFCGCTACCTXYXCRXLHXGKXCKXCHXS 39
        11: 1 1111 1 1 1 1:
Db      1016  CTATTCAGCGCTAACCA--TC-----GGCTTCTACCA 1043

RESULT 10
US-09-252-991A-29199
: Sequence 29199, Application US/09252991A
: Patent No 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 29199
: LENGTH: 281
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29199

```

Query Match	27.1%;	Score 54;	DB 4;	Length 281;
Best Local	Similarity 26.5%;	Pred NO. 11;		
Matches 9;	Conservative 5;	Mismatches 10;	Indels 10;	Gaps 2

Qy 5 CTSTYXCKKFCGTA^{CT}XXYXCRLHXGXKXCXCHC 38
| : | : | : : | : |
Db 232 CSSFCCSSFC-----CSSFC-----SSCSTSYC 255

```

RESULT 11
US-09-627-650B-7
Sequence 7, Application US/09627650B
Patent No. 6406872
GENERAL INFORMATION:
APPLICANT: Bamber, Bruce
APPLICANT: Jorgensen, Erik
TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
FILE REFERENCE: 21101.0009U3
CURRENT APPLICATION NUMBER: US/09/627,650B
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 09/436,063
PRIOR FILING DATE: 1999-11-08
PRIOR APPLICATION NUMBER: 60/107,727
PRIOR FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 2508
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-627-650B-7

```

Query Match	26.9%	Score 53.5	DB 4	Length 2508
Best Local Similarity	30.6%	Pred. No. 76		
Matches 11; Conservative		Mismatches 19		Indels 5; Gaps 2

0y 5 CTSYXCKKFC-GTAXCTYXXCPRLHXGKXCXCHCS 39
||| | | | | | | | :
Db 1007 CGTGCTCACAATGACAACTC----TGATCACTACA 1038

RESULT 12
US-09-436-063C-7
; Sequence 7, Application US/09436063C
; Patent No. 6407210

```

1  GENERAL INFORMATION:
2  APPLICANT: Bamberg, Bruce
3  APPLICANT: Jorgensen, Erik
4  TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
5  TITLE OF INVENTION: Methods Related Thereto
6  FILE REFERENCE: P-1095corrected
7  CURRENT APPLICATION NUMBER: US/09/436,063C
8  CURRENT FILING DATE: 1999-11-08
9  PRIOR APPLICATION NUMBER: 60/107727
10 PRIOR FILING DATE: 1998-11-09
11 NUMBER OF SEQ ID NOS: 18
12 SOFTWARE: PatentIn Ver. 2.1
13 SEQ ID NO 7
14 LENGTH: 2508
15 TYPE: PRT
16 ORGANISM: Caenorhabditis elegans
17 US-09-436-063C-7

```

Query Match	26.9%	Score	53.5	DB	4	Length	2508
Best Local Similarity	30.6%	Pred.	No. 76				
Matches	11	Conservative	1	Mismatches	19	Indels	5
						Gaps	2

QY 5 CTSTYXCKAFEC-GTAXCTYYXCRRXLHGXKXCXCXHC 39
||| | | | | | :
Db 1007 CHTGTGCACAAATGACAACTC---TGATCACTACA 1038

```

RESULT 13
US-09-627-650B-3
Sequence 3, Application US/09627650B
Patent No. 6406872
GENERAL INFORMATION:
APPLICANT: Bamber, Bruce
APPLICANT: Jorgensen, Erik
TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
FILE OF INVENTION: Methods Related Thereto
FILE REFERENCE: 21101.0009U3
CURRENT APPLICATION NUMBER: US/09/627,650B
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 09/436,063
PRIOR FILING DATE: 1999-11-08
PRIOR APPLICATION NUMBER: 60/107,727
PRIOR FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 2544
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-627-650B-3

```

Query Match	26.9%	Score	53.5	DB	4	Length	2544
Best Local Similarity	30.6%	Pred. NO.	77				
Matches	11	Conservative	1	Mismatches	19	Indels	5
						Gaps	2

Oy 5 CTSTYXCAKFC-GTAXCTYYXCPLHXGKXCXCXHC 39
||| | | | | :
Db 1007 CTTGTCTCACAATGCACACTC---TGATCACTACA 1038

```

RESULT 14
US-09-436-063C-3
; Sequence 3, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamberg, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; TITLE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727

```

; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2544
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-3

Query Match 26.9%; Score 53.5; DB 4; Length 2544;
Best Local Similarity 30.6%; Pred. No. 77;
Matches 11; Conservative 1; Mismatches 19; Indels 5; Gaps 2;

QY 5 CTSTYXCKKFC-GTAXCTYYXCRXLHXGKXCXCHCS 39
DB 1007 CTGTCTCACAATGACAACATC---TGATCACTACA 1038

RESULT 15
US-09-627-650B-9
; Sequence 9, Application US/09627/650B
; Patent No. 6406872
; GENERAL INFORMATION:

; APPLICANT: Bamberg, Bruce
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-9

Query Match 26.9%; Score 53.5; DB 4; Length 2601;
Best Local Similarity 30.6%; Pred. No. 78;
Matches 11; Conservative 1; Mismatches 19; Indels 5; Gaps 2;

QY 5 CTSTYXCKKFC-GTAXCTYYXCRXLHXGKXCXCHCS 39
DB 1007 CTGTCTCACAATGACAACATC---TGATCACTACA 1038

Search completed: September 3, 2003, 11:03:23
Job time : 20.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:54:26 ; Search time 60.5 Seconds
(without alignments)
170.613 Million cell updates/sec

Title: US-10-030-231-6
Perfect score: 252
Sequence: 1 HSHACTSYWCGKFCGTCSTHYLCRVLPKMKACVHCGR 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mmc:.*
8: sp_organelle:.*
9: sp_plant:.*
10: sp_proteus:.*
11: sp_virus:.*
12: sp_vertebrate:.*
13: sp_unclassified:.*
14: sp_virus:.*
15: sp_bacteriap:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70.5	28.0	479	4	Q96JVI
2	65	23.8	411	13	P70046
3	65	23.8	460	11	Q99PC6
4	65	25.8	460	11	Q91WN8
5	64	25.4	188	11	Q920U6
6	64	25.4	411	13	P70043
7	63	25.0	143	4	Q9P0D0
8	63	25.0	721	5	Q95YGO
9	63	25.0	1170	10	Q91FZ7
10	62	24.6	438	5	Q81497
11	61	24.2	109	17	Q975R2
12	60	23.8	129	17	Q97ZY4
13	60	23.8	1119	5	Q18034
14	60	23.8	2214	5	Q20219
15	59.5	23.6	1308	5	Q9GPM8
16	59	23.4	573	13	P70045

17	59	23.4	573	13	Q9PSY7	Q9PSY7 xenopus lae
18	59	23.4	856	12	Q80UT7	Q80UT7 infectious
19	58.5	23.2	172	11	Q8C6L7	Q8C6L7 mus musculu
20	58.5	23.2	505	16	Q91037	Q91037 streptomyc
21	58.5	23.2	2515	5	Q24531	Q24531 drosophila
22	58.5	23.2	2731	5	Q9VNU6	Q9VNU6 drosophila
23	58.5	23.2	2731	5	Q61307	Q61307 drosophila
24	58.5	23.2	2731	5	Q18366	Q18366 drosophila
25	58	23.0	241	10	Q8H8G4	Q8H8G4 oryza sativ
26	58	23.0	420	5	P91776	P91776 pacifastacu
27	58	23.0	1895	10	Q9S267	Q9S267 arabidopsis
28	57.5	22.8	88	10	Q9ZTWM	Q9ZTWM petunia hyb
29	57.5	22.8	159	4	Q9BYQ3	Q9BYQ3 homo sapien
30	57.5	22.8	163	12	Q8VB89	Q8VB89 white spot
31	57.5	22.8	337	5	Q8WYJ9	Q8WYJ9 caenorhabdi
32	57.5	22.8	483	6	Q22423	Q22423 caenorhabdi
33	57.5	22.8	732	6	Q28484	Q28484 macaca fasc
34	57	22.6	247	3	Q12531	Q12531 saccharomyc
35	57	22.6	3191	5	Q01335	Q01335 caenorhabdi
36	56.5	22.4	205	15	Q88288	Q88288 snakehead r
37	56.5	22.4	1512	13	Q9DE36	Q9DE36 brachydanio
38	56	22.2	172	12	P88972	P88972 kyzylagach
39	56	22.2	251	5	Q24774	Q24774 enchytraeus
40	56	22.2	437	5	Q81RT5	Q81RT5 drosophila
41	56	22.2	456	6	Q9YTR0	Q9YTR0 canis famill
42	56	22.2	612	4	Q9NVJ6	Q9NVJ6 homo sapien
43	56	22.2	840	5	Q9VZF2	Q9VZF2 drosophila
44	56	22.2	1329	4	Q9HCE3	Q9HCE3 homo sapien
45	56	22.2	4823	13	Q93321	Q93321 fugu rubrip

ALIGNMENTS

RESULT 1
ID Q96JVI PRELIMINARY: PRT: 479 AA.
AC Q96JVI: 01-DEC-2001 (TREMUREL. 19, Created)
DT 01-DEC-2001 (TREMUREL. 19, Last sequence update)
DT 01-MAR-2003 (TREMUREL. 23, Last annotation update)
DE Hypothetical protein F114959.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.,
RT "NEBO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
EMBL AK027865; BAB55416.1;
DR InterPro: IPR007087; Znf_C2H2.
DR InterPro: IPR007086; Znf_C2H2_sub.
DR InterPro: IPR006025; Zn_MTPetide.
DR Pfam: PF000096; zfc_C2H2; 6.
DR PRINTS: PR00048; ZINC_FINGER.
DR SMART: SM00355; Znf_C2H2; 6.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 5.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Metal-binding; Nuclear protein; zinc;
KW zinc-finger.
SQ
SEQUENCE 479 AA: 55354 MW: B180AE85F132319 CR664;
Query Match. 28.0%; Score 70.5; DB 4; Length 479;
Best Local Similarity 47.4%; Pred. No. 0.33;

	Matches	18;	Conservative	2;	Mismatches	13;	Indels	5;	Gaps	4.			
QY		7	SYMGKFCGTA--SCHNYLR--VHPPGM-CACVHCGR	40	:								:
Db		121	AVECTK-CGAKFPCDSTLCRHEVTHSGKKPCECKCGK	157									

RESULT 2
P70046
ID P70046 PRELIMINARY; PRT; 411 AA

DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE XP_0141.
OS *Xenopus laevis* (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; *Xenopus*.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97321550; PubMed=9178260;
RA Bellefroid E., Bourguignon C., Bouwmeester T., Rausch O., Blumberg B.,
RA Pieler T.;
RT "Transcription regulation and alternative splicing of an early zygotic
RT gene encoding two structurally distinct zinc finger proteins in
RT *Xenopus laevis*.";
RL Mech. Dev. 63:99-108(1997).
DR EMBL: U67077; AAB72012.1; -.
DR HSPR: P25490; 10BD.
DR InterPro: IPR007087; ZnF_C2H2.
DR Pfam: PF00096; zf_C2H2; 12.
DR ProDom: PD000003; ZnF_C2H2; 3.
DR SMART: SM00355; ZnF_C2H2; 12.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 12.
KW Metal-binding; Zinc; Zinc-finger;
SO SEQUENCE 411 AA; 46530 MW; E7309EBE9FB9704 CRC64;

Query Match	25.8%	Score 65;	DB 13;	Length 411;
Best Local Similarity	42.9%	Pred. NO. 1.5;		
Matches 18; Conservative	3;	Mismatches 13;	Indels 8;	Gaps 4;

QY 2 SHACTSYMCGKFCGTAS--CTHYLCRVLHPG-KMCACVHCSR 40
 ||| : ||| | | | | | | :
 Db 67 SFACRE--CGQGFRKRSKLTHTFLC--HTGEKPCVCVHCSK 103

RESULT 3	
Q99PC6	
ID	Q99PC6
NO	000006
PRELIMINARY;	
PRT;	460 AA

DT	01-JUN-2001	(TREMblrel, 17, last sequence update)
DT	01-MAR-2003	(TREMblrel, 23, last annotation update)
DE	Anticoagulant protein C.	

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus
OX NCBI_TaxID=10090;

RP	SEQUENCE FROM
RC	STRAIN=C57BL/

Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases

EMBL: AF318182: AAK07918.1: -

DR HSSP; P04070; 1AUT

```
DR  MGD; MGI:97771; Proc.  
DR  InterPro; IPR000152; Asx_hydroxyl.
```

DR InterPro: IPR0013814, Chymotrypsin.
DR InterPro: IPR001881, EGF_Cat.
DR InterPro: IPR006209, EGF_Like.
DR InterPro: IPR002383, GLA_blood.
DR InterPro: IPR001254, Ser_protease_Try.
DR InterPro: IPR000294, Vitk_dep_GLA.
DR RefSeq: NP_000008.2, Vitk.

DR PFAM: PF00594; gls: 1.
DR PFAM: PF00089; trypsin: 1.
DR PRINTS: PR00722; CHYMOTRYPSIN
DR PRINTS: PR00001; CHYMOTRYPSIN

```
DR SMART; SM000179; EGF_CA; 1.
DR SMART; SM000699; GLA; 1.
DR SMART; SM000200; TYPY_SPE; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1
DR PROSITE; PS00033; PCN_1; 1
```

DR PROSITE: PS00240 TRYPSIN_DOM. 1
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00017; EGF_3; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Hydrolyase; Protease; Serine protease
SQ SEQUENCE 460 AA; 51784 MW; 0293BC25ED3FE016 CRC64

Query Match	25.8%;	Score 65;	DB 11;	Length 460;
Best Local Similarity	33.3%;	Pred. No. 1.6;		
Matches 15; Conservative	2;	Mismatches 16;	Indels 12;	Gaps 2;

QY 2 SHACTSYNGKFC-----GTASCTHYLCRVLHPGKMCAC 35
| : | | | | | | | : | | |
Db 118 SCSCDKGEGKFCQQLRFQDCRVNNGCLHY-CLEESNGRRAC 161

RESULT 4	PRELIMINARY;	PRT;	460 AA
Q91WN8			
ID			
Q91WN8			
NO			
COLTWO			

DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update),
DE Similar to protein C.

0C Mus musculus (Mouse).
0S Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
0C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
0X NCBI_TaxId=10090;

RP TISSUE=Liver;
RC SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases

DR EMEL; BC013896; AAH13896.1; -.
DR HSSP; P00761; IANI.
DR MGD; MG1:97771; PLOC.
InterPro: IPR000152. *See below*

DR	InterPro: IPR001881;	EGF_Ca.
DR	InterPro: IPR006209;	EGF_like.
DR	InterPro: IPR002383;	GAL_blood.
DR	InterPro: IPR001254;	Gal_protect.

DR interlepio; LPR
DR Pfam; PF00008
DR Pfam: PF00594

DR PRINTS; PRO0722; CHYMOTRYPSIN
DR PRINTS; PRO0001; GLABLOOD.

DR SMART; SM00179; EGF_CA; 1.

DR SMART; SM00069; GLA; 1.
DR SMART: SM00020; TYPD SDC;

```

DR      CENNA; 3500020; 11yp_spc; 1.
DR      PROSITE; PS00010; ASX_HYDROXYL; 1
DR      PROSITE; PS00023; EGF_1; 1

```

DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS00011; GLU-CARBOXYLATION; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR EGF-like domain; Hydrolase; Protease; Serine protease.
 KW SEQUENCE 460 AA; 51818 MW; 0117F26E68FC274 CRC64;

Query Match 25.8%; Score 65; DB 11; Length 460;
 Best Local Similarity 33.3%; Pred. No. 1.6;
 Matches 15; Conservative 2; Mismatches 16; Indels 12; Gaps 2;

OY 2 SHACTSYWCGKFC-----GTASCTHYLCRVLPKMKAC 35
 DB 118 SCSDKMGKCKFCQGLRFDCCRVNNGCLHY-CLEESNGRCAC 161

RESULT 5

O9Z0U6 PRELIMINARY; PRT; 188 AA.
 AC O9Z0U6;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Late gestation lung protein 1.
 GN LGI1.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-WISTAR; TISSUE=Lung;
 RX MEDLINE-99292450; PubMed-10362728;
 RA Kaplan F., Ledoux P., Kassamali F.Q., Gagnon S., Post M., Koehler D.,
 RA Delmaling J., Swezey N.B.;
 RT "A novel developmentally regulated gene in lung mesenchyme: homology
 RT to a tumor derived trypsin inhibitor."
 RL Am. J. Physiol. 276:LI027-LI036(1999).

DR EMBL; AF109674; AAD16986.1; -.
 DR HSSP; P04284; ICFE.
 DR InterPro; IPR001283; Allrgn_V5/TpXl.
 DR Pfam; PF00188; SCP; 1.
 DR PRINTS; PR00837; V5TFPLIKE.
 DR PRODOM; PD000542; Allrgn_V5/TpXl; 1.
 DR SMART; SM00198; SCP; 1.
 DR PROSITE; PS01010; SCP_AGS_PRL-SC7_2; 1.

SO SEQUENCE 188 AA; 21907 MW; 5B6CC77515360BCE CRC64;

Query Match 25.4%; Score 64; DB 11; Length 188;
 Best Local Similarity 36.6%; Pred. No. 1;
 Matches 15; Conservative 2; Mismatches 22; Indels 2; Gaps 2;

OY 1 HSHACTSYWCGKFCGTASCTHYLCRVLPKMKAC-VHCSR 40
 DB 82 YPHECNP-WCPERCSCAMCTHYTOMWATTNKICGAVHTCR 121

RESULT 6

P70043 PRELIMINARY; PRT; 411 AA.

AC P70043;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Zinc finger protein XFDL 141.

OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Bourguignon C., Bellefroid E., Bouwmeester T., Rausch O., Pieler T.;
 RT "XFDL 141."
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U65897; AAB07010.1; -.
 DR HSSP; P25490; I0BD.
 DR InterPro; IPR007087; ZnF_C2H2.
 DR Pfam; PF00096; zf-C2H2; 12.
 DR PRODOM; PD000003; ZnF_C2H2; 3.
 DR SMART; SM00355; ZnF_C2H2; 12.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
 DR PROSITE; PS01571; ZINC_FINGER_C2H2_2; 12.
 KW Metal-binding; Zinc; Zinc-finger.

SO SEQUENCE 411 AA; 46615 MW; 0F3984747406A23D CRC64;

Query Match 25.4%; Score 64; DB 13; Length 411;
 Best Local Similarity 54.5%; Pred. No. 2;
 Matches 12; Conservative 2; Mismatches 4; Indels 4; Gaps 2;

OY 20 THYLCRVLPK-KMKACVHCSR 40
 DB 85 THFLC---HTGEKPCVCHCK 103

RESULT 7

O9P0D0 PRELIMINARY; PRT; 143 AA.

AC O9P0D0;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HSPC106 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=umbilical cord blood;

RA Zhang Q.H., Ye M., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
 RA Fan H.T., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
 RT "Human partial CDS cloned from cd34+ stem cells."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF161369; AAF28929.1; -.
 FT NON TER 1
 SO SEQUENCE 143 AA; 15676 MW; 52B1ED092E90E68F CRC64;

Query Match 25.0%; Score 63; DB 4; Length 143;
 Best Local Similarity 52.0%; Pred. No. 1.1;
 Matches 13; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

OY 11 GKFCGTASCTHYLCRVLPKMKAC 35
 DB 19 GKFGASAFWDSL--ALHPGRCLPC 41

RESULT 8

O95YG0 PRELIMINARY; PRT; 721 AA.

AC O95YG0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Delta.

GN CS-DELTA.
 OS Ciona savignyi.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Clonidae; Clona.
 OX NCBI_TaxID=51511;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Imai K.S., Satoh N., Satou Y.;
 RT "Ciona savignyi protein.";

SECRETOR, R. L. Submitted (JUL-1997) to the EMBL/GenBank/DBD databases.
[3]
SEQUENCE FROM N.A.
ECKER, J. R. Submitted (AUG-1998) to the EMBL/GenBank/DBD databases.
[4]
SEQUENCE FROM N.A.
SHKIN, P. P., Buehler E. E., Dunn P. P., Feng J. J., Kim C. C., Li Y. Y.,
Walker M. A., Altai H. H., Arano R. R., Conn L. L., Conway A. A. B.,
Gonzalez A. A., Hansen N. N. F., Huizar L. L., Kremenechka I. I.,
Leng C. C., Li J. J., Liu S. S., Luros S. S., Rowley D. D., Schwartz J. J.,
Toriumi M. M., Vysotskaya V. V., Yu G. G., Davis R. R. W.,
Federpiel N. N. A., Theologis A. A., Ecker J. J. R. Submitted (DEC-1998) to the EMBL/GenBank/DBD databases.
[5]
SEQUENCE FROM N.A.
ECKER, J. R. Submitted (MAY-2000) to the EMBL/GenBank/DBD databases.
[6]

RESULT 11	
0975R2	
IID	0975R2
AC	PRELIMINARY;
DT	0975R2;
DDT	01-DEC-2001 (TREMblrel, 19, Created)
DDT	01-DEC-2001 (TREMblrel, 19, Last sequence update)
DDT	01-DEC-2001 (TREMblrel, 19, Last annotation update)
IDE	Hypothetical protein ST0359.
GN	Sulfolobus tokodaii.
OS	Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC	Sulfolobus.
NCBI_TaxID	111955;
RRN	[1]
RRP	SEQUENCE FROM N.A.
RC	STRAIN=JCM 10545 / 7;

```

RX PubMed-11572479;
RA Kavaradaayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamaya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yamagiti M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000982; BAB5338.1;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 109 AA; 12721 MW; EAECAAE34506D364 CRC64;

Query Match 24.2%; Score 61; DB 17; Length 109;
Best Local Similarity 42.9%; Pred. No. 1.6;
Matches 12; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

QY 9 WCGKFCGTASCTHYLC--RVLHPGKMA 34
Db 40 WVGDCITGDCQYAVCEKRALLPGNKCA 67

RESULT 12
Q97ZY4 PRELIMINARY; PRT; 129 AA.
ID Q97ZY4;
AC Q97ZY4;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Hypothetical protein SS00438.
GN SS00438.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
CX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-21332296; PubMed-11427726;
RA Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Weiner C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006676; AAK40763.1;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 129 AA; 15036 MW; D1FFBF263622DB69 CRC64;

Query Match 23.8%; Score 60; DB 17; Length 129;
Best Local Similarity 42.9%; Pred. No. 2.4;
Matches 12; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

QY 9 WCGKFCGTASCTHYLC--RVLHPGKMA 34
Db 60 WVGDCITGDCQYAVCEKRALLPGNKCA 87

RESULT 13
Q18034 PRELIMINARY; PRT; 1119 AA.
ID Q18034;
AC Q18034;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 125.1 kDa protein.
GN C16A3.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```

```

OC Rhabditidae; Peloderinae; Caenorhabditis.
RX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX Favello A.;
RT "The sequence of C. elegans cosmid C16A3.";
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX Waterston R.;
RT "Direct submission.";
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: U41534; AAB47395.1;
DR Wormpep; C16A3.7; CE04007.
DR InterPro; IPR001374; R3H.
DR InterPro; IPR000967; znf_NFX1.
DR InterPro; IPR001841; znf_Ring.
DR Pfam; PF01422; ZF-NFX1; 9.
DR SMART; SM00393; R3H; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00438; znf_NFX; 9.
DR PROSITE; PS50089; ZF_RING_2; 1.
KM Hypothetical protein_125147 MW; 067CF6DF66C45DFE CRC64;
SQ SEQUENCE 1119 AA; 125147 MW;

Query Match 23.8%; Score 60; DB 5; Length 1119;
Best Local Similarity 37.5%; Pred. No. 16;
Matches 12; Conservative 3; Mismatches 11; Indels 6; Gaps 2;

QY 10 CGKFCGTA---SCTHYLCRVLHPGKMACVHC 38
Db 318 CGETCGARKFCGPHCTELCHPG---PCIEC 346

RESULT 14
Q20219 PRELIMINARY; PRT; 2214 AA.
ID Q20219;
AC Q20219;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 243.4 kDa protein.
GN F40F4.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX Wilson R.;
RT "The sequence of C. elegans cosmid F40F4.";
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
RN [3]

```

RESULT 15		PRELIMINARY:		PRT: 1308 AA.	
ID	09GPM8				
AC	09GPM8				
DT	01-MAR-2001 (TREMBLrel, 16, Created)				
DT	01-MAR-2001 (TREMBLrel, 16, Last sequence update)				
DT	01-MAR-2003 (TREMBLrel, 23, Last annotation update)				
DE	Notch-like transmembrane receptor.				
GN	GLP-1.				
OS	Caenorhabditis remanei.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;				
OC	Rhabditidae; Pelodierinae; Caenorhabditis.				
OX	NCBI_TaxID=31234;				
RP	[1]				
RN	SEQUENCE FROM N.A.				
RC	STRAIN-SBI46;				
RX	MEDLINE=21100341; PubMed=11156985;				
RA	Rudel D., Kimble J.;				
RT	"Conservation of glp-1 regulation and function in nematodes.";				
RL	Genetics 157:639-654(2001).				
DR	EMBL: AF315556; AAG49318.1;				
DR	HSSP: P00740; IEDM.				
DR	InterPro: IPR002110; ANK.				
DR	InterPro: IPR000152; ASX_HydroxyL.				
DR	InterPro: IPR000742; EGF_2.				
DR	InterPro: IPR001881; EGF_Ca.				
DR	InterPro: IPR006209; EGF_like.				
DR	InterPro: IPR000800; Notch.				
DR	Pfam: PF00023; ank; 6.				
DR	Pfam: PF00008; EGF; 10.				
DR	Pfam: PF00066; notch; 3.				
DR	PRINTS; PR01452; NOTCH.				
DR	SMART; SM00248; ANK; 4.				
DR	SMART; SM00179; EGF_CA; 1.				
DR	SMART; SM00004; NL; 3.				
DR	PROSITE; PSS0088; ANK_REPEAT; 3.				
DR	PROSITE; PSS0297; ANK_REPEAT_REGION; 1.				
DR	PROSITE; PS00010; ASX_HYDROXYL; 2.				

Search completed: September 3, 2003, 11:01:43
Job time : 62.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:55:11 ; Search time 25 Seconds

(without alignments)
153.870 Million cell updates/sec

Title: US-10-030-231-6

Perfect score: 252
Sequence: 1 HSHACTSYMGKFCGTCSTHYLCRVLPGRKACVHCGR 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR_76.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	25.8	461	1 JX0210	protein C (activat
2	63	25.0	1170	2 C96599	protein P20N2.14
3	60	23.8	129	2 D90188	hypothetical prote
4	60	23.8	1119	2 A88481	protein C16A3.6
5	60	23.8	2214	2 T16305	hypothetical prote
6	59	23.4	71	1 N2NUJC	long neurotoxin 1
7	59	23.4	111	2 S06569	finger protein (cl
8	58.5	23.2	212	2 T05936	agglutinin isolect
9	58.5	23.2	963	2 S62935	hypothetical prote
10	58.5	23.2	2406	2 A54148	odx protein - frui
11	58.5	23.2	2515	2 S47008	tenascin-like prot
12	58.5	23.2	5147	1 IJFFPM	cadherin-related t
13	58	23.0	1895	2 T06609	disease resistance
14	58	23.0	3075	2 S14458	laminin alpha-1 ch
15	57.5	22.8	483	2 T24856	hypothetical prote
16	57.5	22.8	520	2 G88846	protein T12A7.2
17	57.5	22.8	732	2 I52361	testicular metallo
18	57	22.6	247	2 S57548	hypothetical prote
19	57	22.6	3191	2 T22945	hypothetical prote
20	56.5	22.4	186	2 A28401	agglutinin isolect
21	56.5	22.4	410	2 S15163	probable transpos
22	56	22.2	72	1 N2NUJA	long neurotoxin 1
23	56	22.2	182	2 A36686	ultra-high-sulfur
24	56	22.2	251	2 A55035	cysteine-rich prot
25	56	22.2	372	1 S28296	hypothetical prote
26	56	22.2	461	1 KXHU	protein C (activat
27	56	22.2	461	1 S18994	protein C (activat
28	55.5	22.0	71	1 N2NU2W	long neurotoxin 2
29	55.5	22.0	157	2 S58000	probable olfactory

30	55.5	22.0	197	2 I46413	keratin KAP5.5 - s
31	55	21.8	148	2 I36931	lysozyme (EC 3.2.1
32	55	21.8	379	2 A35669	gene CYP61 protein
33	55	21.8	551	2 S38086	hypothetical prote
34	54.5	21.6	352	2 T23464	hypothetical prote
35	54.5	21.6	558	2 JC5878	plasma hyaluronan-
36	54	21.4	130	1 LZBA	lysozyme (EC 3.2.1
37	54	21.4	191	2 I46412	keratin KAP5.4 - s
38	54	21.4	272	2 T15351	hypothetical prote
39	54	21.4	559	2 G71327	probable apolipopr
40	54	21.4	602	2 F71617	SERA antigen/papai
41	54	21.4	1162	2 T49191	hypothetical prote
42	54	21.4	2813	1 VWHU	hypothetical prote
43	53.5	21.2	729	2 T23972	von Willebrand fac
44	53.5	21.2	1106	2 T13938	hypothetical prote
45	53.5	21.2	1106	2 T44598	gene shuttle craft

ALIGNMENTS

RESULT 1

JX0210
protein C (activated) (EC 3.4.21.69) precursor - mouse
N:Alternate names: vitamin K-dependent serine proteinase
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JX0210
R:Tada, N.; Sato, M.; Tsujimura, A.; Iwase, R.; Hashimoto-Gotoh, T.
J. Biochem. 111, 491-495, 1992
A:Title: Isolation and characterization of a mouse protein C cDNA.
A:Reference number: JX0210; MUID:92216897; PMID:1618739
A:Accession: JX0210
A:Molecule type: mRNA
A:Residues: 1-461 <TAD>
A:Cross-references: GB:DI0445; NID:9220385; PIDN:BAA01235.1; PID:9220386
A:Experimental source: liver
C:Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that
s.
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
F:1-33/Domain: signal sequence #status predicted <SIG>
F:27-85/Domain: Gla domain homology <GLA>
F:34-41/Domain: propeptide #status predicted <PRO>
F:42-196/Domain: propeptide #status predicted <PRO>
F:42-196/Domain: light chain #status predicted <PCL>
F:91-130/Domain: EGF homology <EG1>
F:139-174/Domain: heavy chain #status predicted <PCH>
F:199-461/Domain: activation peptide #status predicted <ACT>
F:212-461/Domain: vitamin K-dependent serine proteinase #status predicted <VIT>
F:212-445/Domain: trypsin homology <TRY>
F:47/48/55/57/60/61/66/67/70/76/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F:112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F:121-130/139-150/146-159/161-174/182-191/238-254/373-387/398-426/Disulfide bonds: #s
F:124/299/355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:253/299/402/Active site: His, Asp, Ser #status predicted

Query Match 25.8%; Score 65; DB 1; Length 461;
Best Local Similarity 33.3%; Pred. No. 6.8;
Matches 15; Conservative 2; Mismatches 16; Indels 12; Gaps 2;

OY 2 SHACTSYMGKFC-----GTASCTHYLCRVLPGRKAC 35
DB 118 SCSCDKGWGKFCQDELPRDCCRVNNGGLHY-CLESNGKRCAC 161

RESULT 2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96599

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marshall, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C96599
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1170 <STO>
 A:Cross-references: GB:AE005173; NID:98778496; PIDN:AAF79504.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F20W2.14
 A:Map position: 1

Query Match 25.0%; Score 63; DB 2; Length 1170;
 Best Local Similarity 31.0%; Pred. No. 22;
 Matches 13; Conservative 5; Mismatches 16; Indels 8; Gaps 1;

OY 3 HACTSYWGRKFCGASCTHYLCRVLHP-----GKMKACV 36
 DB 326 HAITGWMGVKPGVSGMHEYSKAYRVPVSSPGVSSSENGCL 367

RESULT 3

D90188
 hypothetical protein SS00438 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: D90188

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan, Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.

arrett, R.A.; Ragun, M.A.; Senner, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: D90188

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-129 <CUR>

A:Cross-references: GB:AE006641; NID:g13813591; PIDN:AAK40763.1; GSPDB:GN00155

C:Genetics:

A:Gene: SS00438

Query Match 23.8%; Score 60; DB 2; Length 129;
 Best Local Similarity 42.9%; Pred. No. 9.4;
 Matches 12; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

OY 9 WCGKFCGTASCTHYLC--RYLHPGKMC 34
 DB 60 WVGDVCTGECQYAYCEKRALPGKCA 87

RESULT 4

A88481
 protein C16A3.6 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: A88481

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: A88481

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1119 <STO>
 A:Cross-references: GB:chr_III; PIDN:AB47595.1; PID:g1109825; GSPDB:GN00021; CESP:C1
 C:Genetics:
 A:Gene: C16A3.6
 A:Map position: 3

Query Match 23.8%; Score 60; DB 2; Length 1119;
 Best Local Similarity 37.5%; Pred. No. 44;
 Matches 12; Conservative 3; Mismatches 11; Indels 6; Gaps 2;

OY 10 CGKFCGTA---SCTHYLCRVLHPGKMCACVHC 38
 DB 318 CGTCGAKKFCGCPHCTELCHPG---PCIEC 346

RESULT 5

T16305
 hypothetical protein F40F4.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16305

R:Wilson, R.
 submitted to the EMBL Data Library, November 1995

A:Description: The sequence of *C. elegans* cosmid F40F4.

A:Reference number: Z18493

A:Accession: T16305

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2214 <MIL>
 A:Cross-references: EMBL:U040420; NID:g1065513; PID:g1065514; PIDN:AAA81430.1; CESP:F4

C:Genetics:

A:Gene: CESP:F40F4.6

A:Introns: 57/3; 95/1; 302/3; 323/1; 380/1; 404/1; 468/1; 507/2; 547/3; 595/3; 1532/1

Query Match 23.8%; Score 60; DB 2; Length 2214;
 Best Local Similarity 31.0%; Pred. No. 72;
 Matches 13; Conservative 5; Mismatches 16; Indels 8; Gaps 1;

OY 5 CTSTWGRKFC-----GTASCTHYLCRVLHPGKMCACVHC 38
 DB 856 CTPEFTGYCONIICVNGTASGDHCVCPPGVAGSECMARC 897

RESULT 6

N2NJUC
 long neurotoxin 1 - Cape cobra

N:Alternate names: neurotoxin alpha

C:Species: Naja naja (Cape cobra)
 C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 23-Aug-1996

C:Accession: A01661

R:Botes, D.P.
 J. Biol. Chem. 246, 7383-7391, 1971

A:Title: Snake venom toxins. The amino acid sequences of toxins alpha and beta from N

A:Reference number: A92095; MUID:72068724; PMID:5167022

A:Accession: A01661

A:Molecule type: protein

A:Residues: 1-71 <BOT>

C:Superfamily: snake toxin

C:Keywords: neurotoxin; venom

E:3-20,14-26,30,45-56,57-62/Disulfide bonds: #status experimental

Query Match 23.4%; Score 59; DB 1; Length 71;
 Best Local Similarity 35.6%; Pred. No. 7.9;
 Matches 16; Conservative 1; Mismatches 18; Indels 10; Gaps 4;

OY 3 HAC-TSYWGRKFCG-----TASCTHYLCRVLHPGKMCACVHC 40
 DB 18 HVCYTRKMCNDNFCGMRKRVLDGCA-ATCPKVPKGVNKC--CSR 59

RESULT 7

S06569
 finger protein (clone X1CGF32.1) - African clawed frog (fragment)


```

C:Species: Xenopus laevis (African clawed frog)
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 31-Dec-1993
C:Accession: S06569
R:Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeling, A.; Knoech
J. Mol. Biol. 208, 639-659, 1989
A>Title: Second-order repeats in Xenopus laevis finger proteins.
A:Reference number: S05632; MUID:90040698; PMID:2509712
A:Accession: S06569
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-111 <NIF>
C:Keywords: DNA binding; zinc finger

Query Match      23.4%; Score 59; DB 2; Length 111;
Best Local Similarity 42.9%; Pred. No. 11;
Matches 18; Conservative 2; Mismatches 14; Indels 8; Gaps 4;

OY      2 SHACTSYWCGKFCGTASC--CTHYLCRVLHPG-KMCACVHCSR 40
DB      5 SFDCTE--CGKSFRRSKLKTHTFLC---HTGKPPVCVHCGK 41

RESULT 8
T05936
agglutinin isolectin 1 precursor - barley
C:Species: Hordeum vulgare (barley)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000
C:Accession: T05936
R:Lehner, D.R.; Raikhel, N.V.
Plant Physiol. 91, 124-129, 1989
A>Title: Cloning and characterization of root-specific barley lectin.
A:Reference number: 215461
A:Accession: T05936
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-212 <LEK>
A:Cross-references: EMBL:M29280; NID:g167070; PIDN:AAA32969.1; PID:g167071
C:Superfamily: wheat agglutinin; hevein chitin-binding domain homology
C:Keywords: lectin
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-212/Product: agglutinin isolectin 1 #status predicted <MAT>
F:27-69/Domain: hevein chitin-binding domain homology <HCB>

Query Match      23.2%; Score 58.5; DB 2; Length 212;
Best Local Similarity 32.6%; Pred. No. 19;
Matches 14; Conservative 3; Mismatches 17; Indels 9; Gaps 2;

OY      5 CTSY-----WCGKFCGTASC-THYLCRVLHPGKMCACVHC 38
DB      44 CSQYGYCGMGDYGCGKCGACGYTSKRCGTQAGKTCFNNHC 86

RESULT 9
S62935
hypothetical protein YNL023c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N2812
C:Species: Saccharomyces cerevisiae
C>Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 19-Apr-2002
C:Accession: S62935; S62945
R:Andre, B.; Itagui Housaini, I.; Urrestarazu, L.A.; Vissers, S.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S62920
A:Accession: S62935
A:Molecule type: DNA
A:Residues: 1-965 <AND>
A:Cross-references: EMBL:Z71299; NID:g1301854; PID:e239870; PID:g1301855; MIPS:YNL023c
A:Experimental source: strain S288C
R:Duesterhoeft, A.; Floeth, M.; Fritz, C.; Heuss-Neltzel, D.; Hilbert, H.; Moestl, D.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S62944
A:Accession: S62945
A:Molecule type: DNA
A:Residues: 1-965 <DUE>

A:Cross-references: EMBL:Z71299; NID:g1301854; PID:e239870; PID:g1301855; MIPS:YNL023
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:RAP1
A:Cross-references: SGD:S0004968
A:Map position: 14L

Query Match      23.2%; Score 58.5; DB 2; Length 965;
Best Local Similarity 26.3%; Pred. No. 58;
Matches 10; Conservative 6; Mismatches 15; Indels 7; Gaps 1;

OY      10 CGKFCGTASCTHYLCRVLHPGKMCACV-----HCSR 40
DB      150 CGQTCNASTCMHGCSKICHLGPHPECTRMVEIMCHCGK 187

RESULT 10
A54148
odt protein - fruit fly (Drosophila sp.)
C:Species: Drosophila sp.
C>Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 11-Jan-2000
C:Accession: A54148
R:Levine, A.; Bashan-Ahrend, A.; Budai-Hadrian, O.; Gartenberg, D.; Menasherow, S.; W
Cell 77, 587-598, 1994
A>Title: odt O2: a novel Drosophila pair rule gene.
A:Reference number: A54148; MUID:94243925; PMID:7514504
A:Accession: A54148
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2406 <LEV>
C:Genetics:
A:Gene: FlyBase:Ten-m
A:Cross-references: FlyBase:FBgn0004449
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:514-540/Domain: EGF homology <EGF>
F:610-637/Domain: EGF homology <EGF1>

Query Match      23.2%; Score 58.5; DB 2; Length 2406;
Best Local Similarity 29.8%; Pred. No. 1,1e+02;
Matches 14; Conservative 4; Mismatches 14; Indels 15; Gaps 3;

OY      4 ACTSYWCGKFCGTASC-----THYLCRVLHPGKMCACV-----HCS 39
DB      696 ACDPFWGGEYCWTRLCDVRCNEHGQCK---NGTCICLTGMWNGKHCT 738

RESULT 11
S47008
tenascin-like protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C:Accession: S47008
R:Baumgartner, S.; Martin, D.; Hagios, C.; Chiquet-Ehrismann, R.
EMBO J. 13, 3728-3740, 1994
A>Title: Ten(m), a Drosophila gene related to tenascin, is a new pair-rule gene.
A:Reference number: S47008; MUID:94349920; PMID:8070401
A:Accession: S47008
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2515 <BAU>
A:Cross-references: EMBL:X73154; NID:g510505; PIDN:CAA51678.1; PID:g510506
C:Genetics:
A:Gene: FlyBase:Ten-m
A:Cross-references: FlyBase:FBgn0004449
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:298-324/Domain: EGF homology <EGF>
F:394-421/Domain: EGF homology <EGF1>

Query Match      23.2%; Score 58.5; DB 2; Length 2515;
Best Local Similarity 29.8%; Pred. No. 1,1e+02;
Matches 14; Conservative 4; Mismatches 14; Indels 15; Gaps 3;

OY      4 ACTSYWCGKFCGTASC-----THYLCRVLHPGKMCACV-----HCS 39

```

Db 480 ACDEMGECYCNTRLCVRCNHEGCK-----NGTCLCVTGMNGKHC 522

RESULT 12

JTFPM

cadherin-related tumor suppressor precursor - fruit fly (*Drosophila melanogaster*)C:Species: *Drosophila melanogaster*

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997

C:Accession: A41087; B41087

R:Mahoney, P.A.; Weber, U.; Onofrechuk, P.; Blesmann, H.; Bryant, P.J.; Goodman, C.S.

Cell 67, 853-868, 1991

A>Title: The fat tumor suppressor gene in *Drosophila* encodes a novel member of the cadherin family

A:Reference number: A41087; MUID:92069752; PMID:1959133

A:Accession: A41087

A:Molecule type: mRNA

A:Residues: 143-485;1279-5147 <MAH>

A:Cross-references: GB:M80537

A:Accession: B41087

A:Molecule type: DNA

A:Residues: 1-142;487-1278 <MA2>

A:Cross-references: GB:M80537

A>Note: 1229-Gly and 1233-Ser were also found

C:Genetics:

A:Gene: fat

A:Cross-references: FlyBase:FBgn0001075

C:Superfamily: cadherin-related tumor suppressor; cadherin repeat homology; EGF homology

C:Keywords: calcium binding; cell adhesion; duplication; transmembrane protein

F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-5147/Product: cadherin-related tumor suppressor #status predicted <MAT>

F:36-4588/Domain: extracellular #status predicted <EXT>

F:51-156/Domain: cadherin repeat homology <CR1>

F:159-270/Domain: cadherin repeat homology <CR2>

F:271-382/Domain: cadherin repeat homology <CR3>

F:390-494/Domain: cadherin repeat homology <CR4>

F:497-599/Domain: cadherin repeat homology <CR5>

F:602-708/Domain: cadherin repeat homology <CR6>

F:718-822/Domain: cadherin repeat homology <CR7>

F:831-942/Domain: cadherin repeat homology <CR8>

F:948-1049/Domain: cadherin repeat homology <CR9>

F:1052-1153/Domain: cadherin repeat homology <CR10>

F:1156-1278/Domain: cadherin repeat homology <CR11>

F:1281-1384/Domain: cadherin repeat homology <CR12>

F:1387-1489/Domain: cadherin repeat homology <CR13>

F:1492-1601/Domain: cadherin repeat homology <CR14>

F:1607-1713/Domain: cadherin repeat homology <CR15>

F:1717-1823/Domain: cadherin repeat homology <CR16>

F:1826-1922/Domain: cadherin repeat homology <CR17>

F:1925-2027/Domain: cadherin repeat homology <CR18>

F:2028-2167/Domain: cadherin repeat homology <CR19>

F:2169-2278/Domain: cadherin repeat homology <CR20>

F:2281-2384/Domain: cadherin repeat homology <CR21>

F:2387-2491/Domain: cadherin repeat homology <CR22>

F:2494-2596/Domain: cadherin repeat homology <CR23>

F:2599-2703/Domain: cadherin repeat homology <CR24>

F:2707-2810/Domain: cadherin repeat homology <CR25>

F:2813-2913/Domain: cadherin repeat homology <CR26>

F:2915-3013/Domain: cadherin repeat homology <CR27>

F:3014-3124/Domain: cadherin repeat homology <CR28>

F:3127-3229/Domain: cadherin repeat homology <CR29>

F:3232-3334/Domain: cadherin repeat homology <CR30>

F:3337-3439/Domain: cadherin repeat homology <CR31>

F:3442-3545/Domain: cadherin repeat homology <CR32>

F:3548-3651/Domain: cadherin repeat homology <CR33>

F:3654-3756/Domain: cadherin repeat homology <CR34>

F:3954-4010/Domain: EGF homology <EG1>

F:4017-4048/Domain: EGF homology <EG2>

F:4056-4089/Domain: EGF homology <EG3>

F:4096-4127/Domain: EGF homology <EG4>

F:4584-4609/Domain: transmembrane #status predicted <TM>

F:4610-5147/Domain: intracellular #status predicted <INT>

Query Match 23.2% Score 58.5; DB 1; Length 5147;

Best Local Similarity 32.1%; Pred. No. 1.9e+02;
Matches 18; Conservative 3; Mismatches 14; Indels 21; Gaps 4;

QY 6 TSYWC-----GKFGTAS-----CTH-YLCRYLHPGKACACV-----HCSR 40

Db 4074 SSYFCLCRPGFRGNOCESVSDSCRPNPCLHGLCVSLKFKYKCNCTPGRYGRHCR 4129

RESULT 13

T06609

disease resistance protein homolog F16J13.90 - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999

C:Accession: T06609

R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro

submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15789

A:Accession: T06609

A:Molecule type: DNA

A:Residues: 1-1895 <BEV>

A:Cross-references: EMBL:AL049638; GSPDB:GN00062; ATSP:F16J13.90

A:Experimental source: cultivar Columbia; BAC clone F16J13

C:Genetics:

A:Gene: ATSP:F16J13.90

A:Map position: 4

A:Introns: 67/2; 340/2; 391/3; 607/2; 661/2; 791/2; 1148/3; 1255/3; 1646/2; 1674/3; 1

Query Match 23.0% Score 58; DB 2; Length 1895;

Best Local Similarity 27.6%; Pred. No. 1.1e+02;

Matches 16; Conservative 6; Mismatches 14; Indels 22; Gaps 3;

QY 5 CTSTWCGK-----FC-----GRASCTHY-----LCRYLHPGKACACVHCSR 40

Db 193 CKTYCGKSKASGRLPFCRAHGGKCKSHEDCTGPARGRSGCLMHGGKRCQRENCYK 250

RESULT 14

S14458

laminin alpha-1 chain precursor - human

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 19-Jan-2001

C:Accession: S14458; A34961

R:Haaparanta, T.; Uitto, J.; Ruoslahti, E.; Engvall, E.

Matrix 11, 151-160, 1991

A>Title: Molecular cloning of the cDNA encoding human laminin A chain.

A:Reference number: S14458; MUID:9133420; PMID:1714537

A:Accession: S14458

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-3075 <HAA>

R:Nissinen, M.; Vuolteenaho, R.; Boot-Handford, R.; Kallunki, P.; Tryggvason, K.

Biochem. J. 276, 369-379, 1991

A>Title: Primary structure of the human laminin A chain. Limited expression in human

A:Reference number: S14663; MUID:91264789; PMID:2049067

A:Accession: S14663

A:Molecule type: mRNA

A:Residues: 1-227, 'FE', 230-251, 'MLP', 255-418, 'E', 420-518, 'L', 520-1022, 'V', 1024-1074, 'A:Cross-references: EMBL:X58531; NID:934225; PIDN:CAAA1418.1; PID:934226

R:Olson, D.; Nagayoshi, T.; Fazio, M.; Peltonen, J.; Jaakkola, S.; Sanborn, D.; Sasak

Lab. Invest. 60, 772-782, 1989

A>Title: Human laminin: cloning and sequence analysis of cDNAs encoding A, B1 and B2

A:Reference number: A34961; MUID:89280632; PMID:2733383

A:Accession: A34961

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 'W', 2397-2745, 'L', 2747-3053, 'L', 3055-3072, 'PSP', <OLS>

A>Note: the authors translated the codon AGA for residue 2692 as Pro

C:Genetics:

A:Gene: GDB:LAMAL; LAMA

A:Cross-references: GDB:120135; OMIM:150320

A:Map position: 18p11.32-18p11.22

C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-11k

C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; disulfide

F:1-17/Domain signal sequence #status predicted <SIG>
F:18-3075/Product: laminin alpha-1 chain #status predicted <MAT>
F:18-269/Domain: VI <DOM6>
F:270-516/Domain: V <DOM5>
F:270-334/Domain: laminin-type EGF-like homology <LE1>
F:337-394/Domain: laminin-type EGF-like homology <LE2>
F:397-451/Domain: laminin-type EGF-like homology <LE3>
F:454-500/Domain: laminin-type EGF-like homology <LE4>
F:503-512/Domain: laminin-type EGF-like homology <LE5>
F:517-708/Domain: IVB <DO4B>
F:709-1159/Domain: IIIB <DO3B>
F:709-739/Domain: laminin-type EGF-like homology #status atypical <LE6>
F:742-788/Domain: laminin-type EGF-like homology <LE7>
F:791-846/Domain: laminin-type EGF-like homology <LE8>
F:849-899/Domain: laminin-type EGF-like homology <LE9>
F:902-948/Domain: laminin-type EGF-like homology <LE10>
F:951-995/Domain: laminin-type EGF-like homology <LE11>
F:998-1041/Domain: laminin-type EGF-like homology <LE12>
F:1044-1081/Domain: laminin-type EGF-like homology <LE13>
F:1090-1109/Domain: laminin-type EGF-like homology #status atypical <LE14>
F:1111-1147/Domain: laminin-type EGF-like homology #status atypical <LE15>
F:1150-1159/Domain: laminin-type EGF-like homology #status atypical <LE16>
F:1160-1361/Domain: IVA <DO4A>
F:1362-1553/Domain: IIIA <DO3A>
F:1362-1400/Domain: laminin-type EGF-like homology #status atypical <LE17>
F:1403-1449/Domain: laminin-type EGF-like homology <LE18>
F:1452-1506/Domain: laminin-type EGF-like homology <LE19>
F:1509-1553/Domain: laminin-type EGF-like homology <LE20>
F:1554-2125/Domain: I/II, heptad repeats <DOM2>
F:2116-2120/Region: cell adhesion #status predicted
F:2126-3075/Domain: G <DOMG>
F:2142-2300/Domain: laminin G repeat homology <GL1>
F:2339-2484/Domain: laminin G repeat homology <GL2>
F:2510-2676/Domain: laminin G repeat homology <GL3>
F:2534-2536/Region: cell attachment (R-G-D) motif
F:2739-2888/Domain: laminin G repeat homology <GL4>
F:2916-3073/Domain: laminin G repeat homology <GL5>
F:38-166,555,665,763,801,838,926,952,1045,1407,1579,1596,1678,1689,1698,1717,1804,1894,1929-305/Dsulfide bonds: #status predicted
rate (Asn) (covalent) #status predicted

Query Match	23.0%	Score 58	DB 2	Length 3075
Best Local Similarity	36.8%	Pred. No. 1.5e+02		
Matches	14	Conservative	3	Mismatches 9
				Indels 12
				Gaps 2
QY	1	HSNACTSYMCKRFGCTAS	CTTHYLCLRVLLPQKMCACVH	37
Db	1430	HCDDVCTSGYVYGRVTSASDCA	-----LCCAPH	1456

RESULT 15
T24856
hypothetical protein T12A7.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24856
R:Lennard, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19943
A:Accession: T24856
A:Status: Preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-483 <M1>
A:Cross-references: EMBL:Z73911; PIDN:CAA98142.2; GSPDB:GN00022; CESP:T12A7.2
A:Experimental source: clone T12A7
C:Genetics:
A:Gene: CESP:T12A7.2
A:Map position: 4
A:Introns: 36/3; 71/1; 146/1; 255/2; 287/2; 340/3; 387/1; 449/2

Query Match	22.8%;	Score 57.5;	DB 2;	Length 483;
Best Local Similarity	35.0%;	Pred. NO. 45;		
Matches 14;	Conservative	3;	Mismatches 16;	Indels 7; Gaps 2

```

Ox      5  CTTGACGAGC-----GTASCTHYLCRYLHAGKMCACVHC 38
          |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      123  CTSNMGVGEHCIFRMSGVYKNTSGCELF-GRCTIQKC 161

```

Search completed: September 3, 2003, 11:02:38
Job time : 27.5 secs

Search completed: September 3, 2003, 11:02:38
Job time : 27.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:59:06 : Search time 19.5 Seconds
(without alignments)
281.649 Million cell updates/sec

Title: US-10-030-231-6
Perfect score: 252
Sequence: 1 HSHACTSYWCGKFCGTASTCTHLYLGRVLPGRKACVHCSSR 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues
Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67.5	26.8	2119	US-10-140-472-61	Sequence 61, App1
2	67.5	26.8	2119	US-10-141-761-61	Sequence 61, App1
3	67.5	26.8	2119	US-10-142-885-61	Sequence 61, App1
4	67.5	26.8	2119	US-10-123-155-61	Sequence 61, App1
5	67.5	26.8	2119	US-10-146-731-61	Sequence 61, App1
6	66	26.2	1162	US-10-140-472-185	Sequence 185, App
7	66	26.2	1162	US-10-141-761-185	Sequence 185, App
8	66	26.2	1162	US-10-142-885-185	Sequence 185, App
9	66	26.2	1162	US-10-123-155-185	Sequence 185, App
10	66	26.2	1162	US-10-146-731-185	Sequence 185, App
11	64.5	25.6	1371	US-10-140-472-487	Sequence 487, App
12	64.5	25.6	1371	US-10-141-761-487	Sequence 487, App
13	64.5	25.6	1371	US-10-142-885-487	Sequence 487, App
14	64.5	25.6	1371	US-10-123-155-487	Sequence 487, App
15	64.5	25.6	1371	US-10-146-731-487	Sequence 487, App

16	64.5	25.6	3608	US-10-184-634-433	Sequence 433, App
17	64.5	25.6	3608	US-10-184-634-433	Sequence 433, App
18	64	25.4	188	US-09-808-602-109	Sequence 109, App
19	64	25.4	188	US-09-800-198-95	Sequence 95, App1
20	64	25.4	2095	US-10-184-634-161	Sequence 161, App
21	64	25.4	2095	US-10-184-634-161	Sequence 161, App
22	63.5	25.2	1985	US-10-184-634-27	Sequence 27, App1
23	63.5	25.2	1985	US-10-184-634-27	Sequence 27, App1
24	63	25.0	3240	US-10-184-634-415	Sequence 415, App
25	63	25.0	3240	US-10-184-634-415	Sequence 415, App
26	63	25.0	3244	US-10-184-634-571	Sequence 571, App
27	63	25.0	3244	US-10-184-634-571	Sequence 571, App
28	63	25.0	3690	US-10-184-634-517	Sequence 517, App
29	63	25.0	3690	US-10-184-634-517	Sequence 517, App
30	62.5	24.8	1843	US-10-140-472-309	Sequence 309, App
31	62.5	24.8	1843	US-10-141-761-309	Sequence 309, App
32	62.5	24.8	1843	US-10-142-885-309	Sequence 309, App
33	62.5	24.8	1843	US-10-123-155-309	Sequence 309, App
34	62.5	24.8	1843	US-10-146-731-309	Sequence 309, App
35	62.5	24.8	2651	US-10-184-634-135	Sequence 135, App
36	62.5	24.8	2651	US-10-184-634-135	Sequence 135, App
37	62.5	24.8	3781	US-10-184-634-453	Sequence 453, App
38	62.5	24.8	3781	US-10-184-634-453	Sequence 453, App
39	62	24.6	2031	US-10-140-472-101	Sequence 101, App
40	62	24.6	2031	US-10-141-761-101	Sequence 101, App
41	62	24.6	2031	US-10-142-885-101	Sequence 101, App
42	62	24.6	2031	US-10-123-155-101	Sequence 101, App
43	62	24.6	2031	US-10-146-731-101	Sequence 101, App
44	62	24.6	2089	US-10-140-472-497	Sequence 497, App
45	62	24.6	2089	US-10-141-761-497	Sequence 497, App

ALIGNMENTS

RESULT 1
US-10-140-472-61
; Sequence 61, Application US/10140472
; Publication No. US20030138888A1
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SEQUENCED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330301C168
CURRENT APPLICATION NUMBER: US/10/140,472
PRIORITY FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 61
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-472-61
Query Match 26.8%; Score 67.5; DB 12; Length 2119;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 14; Conservative 3; Mismatches 16; Indels 3; Gaps 2;

OY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
11 : 11 1111:1 111 1:
DB 310 ACAA-CGACCGTAACA--TCATATATACCACTGCT 342

RESULT 2

US-10-141-761-61
Sequence 61, Application US/10141761
Publication No. US20030148432A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C198
CURRENT APPLICATION NUMBER: US/10/141,761
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 61
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo Sapien
US-10-141-761-61

Query Match 26.8%; Score 67.5; DB 12; Length 2119;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 14; Conservative 3; Mismatches 16; Indels 3; Gaps 2;

OY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
11 : 11 1111:1 111 1:
DB 310 ACAA-CGACCGTAACA--TCATATATACCACTGCT 342

RESULT 3

US-10-142-885-61
Sequence 61, Application US/10142885
Publication No. US20030157604A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C248
CURRENT APPLICATION NUMBER: US/10/142,885
CURRENT FILING DATE: 2002-05-10
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 61
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo Sapien
US-10-142-885-61

Query Match 26.8%; Score 67.5; DB 12; Length 2119;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 14; Conservative 3; Mismatches 16; Indels 3; Gaps 2;

OY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
11 : 11 1111:1 111 1:
DB 310 ACAA-CGACCGTAACA--TCATATATACCACTGCT 342

US-10-123-155-61
Sequence 61, Application US/10123155
Publication No. US20030068794A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 61
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-61

Query Match 26.8%; Score 67.5; DB 15; Length 2119;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 14; Conservative 3; Mismatches 16; Indels 3; Gaps 2;

OY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
11 : 11 1111:1 111 1:
DB 310 ACAA-CGACCGTAACA--TCATATATACCACTGCT 342

RESULT 5

US-10-146-731-61
Sequence 61, Application US/10146731
Publication No. US20030129692A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C33
CURRENT FILING DATE: 2002-05-15
Prior Application removed - See File Wrapper or Palm
SEQUENCE OF SEQ ID NOS: 550
SEQ ID NO 61
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo Sapien
US-10-146-731-61

Query Match 26.8%; Score 67.5; DB 16; Length 2119;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 14; Conservative 3; Mismatches 16; Indels 3; Gaps 2;

OY 4 ACTSYWCGKFGCTASTCYHLRCVLRHPGKMCACVHC 39
DB 310 ACAA-CCAGCCTACCA--TCATATACCACTGCT 342

RESULT 6
US-10-140-472-185
Sequence 185, Application US/10140472
Publication No. US20030138888A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C168
CURRENT FILING DATE: 2002-05-16
Prior Application removed - See File Wrapper or Palm
SEQUENCE OF SEQ ID NOS: 550
SEQ ID NO 185
LENGTH: 1162
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-472-185

Query Match 26.2%; Score 66; DB 12; Length 1162;
Best Local Similarity 40.0%; Pred. No. 23;

Matches 14; Conservative 2; Mismatches 19; Indels 0; Gaps 0;
OY 4 ACTSYWCGKFGCTASTCYHLRCVLRHPGKMCACVHC 38
DB 1087 ACGATCGATGTGTATCTGACCTCAAGACCACTC 1121

RESULT 7
US-10-141-761-185
Sequence 185, Application US/10141761
Publication No. US20030148432A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C198
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
SEQUENCE OF SEQ ID NOS: 550
SEQ ID NO 185
LENGTH: 1162
TYPE: DNA
ORGANISM: Homo Sapien
US-10-141-761-185

Query Match 26.2%; Score 66; DB 12; Length 1162;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 14; Conservative 2; Mismatches 19; Indels 0; Gaps 0;

OY 4 ACTSYWCGKFGCTASTCYHLRCVLRHPGKMCACVHC 38
DB 1087 ACGATCGATGTGTATCTGACCTCAAGACCACTC 1121

RESULT 8
US-10-142-885-185
Sequence 185, Application US/10142885
Publication No. US20030157604A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 185
; LENGTH: 1162
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-885-185

Query Match          26.2%; Score 66; DB 12; Length 1162;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 14; Conservative 2; Mismatches 19; Indels 0; Gaps 0;

OY      4 ACTSWGKFGCTASCTHYLCRVLHPGKMCACVHC 38
      11 : 11 111:11 11111
Db      1087 ACGATGCGATTGTATCTGCAGCTCAAGCACCTC 1121

RESULT 9
US-10-123-155-185
; Sequence 185, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 185
; LENGTH: 1162
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-185

Query Match          26.2%; Score 66; DB 15; Length 1162;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 14; Conservative 2; Mismatches 19; Indels 0; Gaps 0;

OY      4 ACTSWGKFGCTASCTHYLCRVLHPGKMCACVHC 38
      11 : 11 111:11 11111
Db      1087 ACGATGCGATTGTATCTGCAGCTCAAGCACCTC 1121

RESULT 10
US-10-146-731-185
; Sequence 185, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
US-10-146-731-185

Query Match          26.2%; Score 66; DB 12; Length 1162;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 14; Conservative 2; Mismatches 19; Indels 0; Gaps 0;

OY      4 ACTSWGKFGCTASCTHYLCRVLHPGKMCACVHC 38
      11 : 11 111:11 11111
Db      1087 ACGATGCGATTGTATCTGCAGCTCAAGCACCTC 1121

RESULT 11
US-10-140-472-487
; Sequence 487, Application US/10140472
; Publication No. US20030138888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 487
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-472-487

Query Match          25.6%; Score 64.5; DB 12; Length 1371;

```


Best Local Similarity 31.8%; Pred. No. 39;
Matches 14; Conservative 2; Mismatches 19; Indels 9; Gaps 1;

OY 5 CTSYWGKFCGTASCTHYL-----CRVLHPGKMCACVHCS 39
DB 137 CTACTCGTGGGTCTTCTTCCTTGCGATACAGCTACAGCT 180

RESULT 12

US-10-141-761-487

; Sequence 487, Application US/10141761
; Publication No. US20030148432A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 487
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-141-761-487

Query Match

Best Local Similarity 25.6%; Score 64.5; DB 12; Length 1371;
Matches 14; Conservative 2; Mismatches 19; Indels 9; Gaps 1;

OY 5 CTSYWGKFCGTASCTHYL-----CRVLHPGKMCACVHCS 39
DB 137 CTACTCGTGGGTCTTCTTCCTTGCGATACAGCTACAGCT 180

RESULT 13

US-10-142-885-487

; Sequence 487, Application US/10142885
; Publication No. US20030157604A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 487
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-885-487

Query Match
Best Local Similarity 31.8%; Pred. No. 39;
Matches 14; Conservative 2; Mismatches 19; Indels 9; Gaps 1;

OY 5 CTSYWGKFCGTASCTHYL-----CRVLHPGKMCACVHCS 39
DB 137 CTACTCGTGGGTCTTCTTCCTTGCGATACAGCTACAGCT 180

RESULT 14

US-10-123-155-487

; Sequence 487, Application US/10123155
; Publication No. US20030068794A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 487
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-487

Query Match
Best Local Similarity 25.6%; Score 64.5; DB 15; Length 1371;
Matches 14; Conservative 2; Mismatches 19; Indels 9; Gaps 1;

OY 5 CTSYWGKFCGTASCTHYL-----CRVLHPGKMCACVHCS 39
DB 137 CTACTCGTGGGTCTTCTTCCTTGCGATACAGCTACAGCT 180

RESULT 15

US-10-146-731-487

; Sequence 487, Application US/10146731
; Publication No. US20030129692A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:57:11 ; Search time 20.5 Seconds
(without alignments)
82.558 Million cell updates/sec

Title: US-10-030-231-6
Perfect score: 252
Sequence: 1 HSHACTSYWCGKFCGCTASTHYLCRVLHPKMKACVHCSR 40

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Dackfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61.5	24.4	1248	3	US-08-882-046-6 Sequence 6, Appl
2	59.5	23.6	1055	3	US-09-214-278-2 Sequence 2, Appl
3	59.5	23.6	1065	2	US-08-400-159-8 Sequence 8, Appl
4	59.5	23.6	1148	3	US-08-882-046-4 Sequence 4, Appl
5	59.5	23.6	1212	3	US-09-214-278-3 Sequence 3, Appl
6	59.5	23.6	1238	3	US-09-214-278-5 Sequence 5, Appl
7	59.5	23.6	1257	3	US-08-611-728A-8 Sequence 8, Appl
8	59.5	23.6	1652	4	US-09-627-650B-1 Sequence 1, Appl
9	59.5	23.6	1652	4	US-09-436-063C-1 Sequence 1, Appl
10	59.5	23.6	1917	4	US-09-627-650B-5 Sequence 5, Appl
11	59.5	23.6	1917	4	US-09-436-063C-5 Sequence 5, Appl
12	59.5	23.6	2508	4	US-09-627-650B-7 Sequence 7, Appl
13	59.5	23.6	2508	4	US-09-436-063C-7 Sequence 7, Appl
14	59.5	23.6	2544	4	US-09-627-650B-3 Sequence 3, Appl
15	59.5	23.6	2544	4	US-09-436-063C-3 Sequence 3, Appl
16	59.5	23.6	2601	4	US-09-627-650B-9 Sequence 9, Appl
17	59.5	23.6	2601	4	US-09-436-063C-9 Sequence 9, Appl
18	58.5	23.0	262	4	US-09-252-991A-22359 Sequence 22359, A
19	58	23.0	3075	2	US-08-460-309-5 Sequence 5, Appl
20	58	23.0	3075	2	US-08-125-077-5 Sequence 5, Appl
21	57	22.2	696	4	US-08-933-711B-7 Sequence 7, Appl
22	56	22.2	109	2	US-08-527-044-2 Sequence 2, Appl
23	56	22.2	109	3	US-09-013-780-2 Sequence 2, Appl
24	56	22.2	409	3	US-09-065-872-2 Sequence 2, Appl
25	56	22.2	409	4	US-09-667-570A-2 Sequence 2, Appl
26	56	22.2	410	3	US-09-065-872-1 Sequence 1, Appl
27	56	22.2	410	4	US-09-667-570A-1 Sequence 1, Appl

28	56	22.2	419	1	US-08-295-411-1 Sequence 1, Appl
29	56	22.2	419	2	US-08-955-471-1 Sequence 1, Appl
30	56	22.2	419	4	US-09-667-570A-3 Sequence 3, Appl
31	56	22.2	419	5	PCR-US92-10242-1 Sequence 1, Appl
32	56	22.2	460	2	US-08-756-506-2 Sequence 2, Appl
33	56	22.2	460	2	US-08-756-506-4 Sequence 2, Appl
34	56	22.2	460	6	5270178-13 Patent No. 5270178
35	56	22.2	460	6	5270178-15 Patent No. 5270178
36	56	22.2	460	6	5270178-16 Patent No. 5270178
37	56	22.2	460	6	5270178-16 Patent No. 5270178
38	56	22.2	461	6	5225537-2 Patent No. 5225537
39	56	22.2	461	6	5270178-2 Patent No. 5270178
40	56	22.2	461	6	5270178-17 Patent No. 5270178
41	56	22.2	461	6	5270178-18 Patent No. 5270178
42	56	22.2	461	6	5460953-3 Patent No. 5460953
43	55.5	22.0	751	2	US-08-836-443-3 Sequence 3, Appl
44	55	21.8	379	1	US-08-468-847B-11 Sequence 11, Appl
45	55	21.8	379	4	US-09-142-569-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-882-046-6
Sequence 6, Application US/08882046
Patent No. 6136952
GENERAL INFORMATION:
APPLICANT: LI, Linheng
APPLICANT: Hood, Leroy
APPLICANT: Krantz, Ian D.
APPLICANT: Spinner, Nancy B.
TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
TITLE OF INVENTION: Nucleic Acids and Methods of Use
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,046
FILING DATE: 25-JUN-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-0W 2637
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1248 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-882-046-6
Query Match 24.4%; Score 61.5; DB 3; Length 1248;
Best Local Similarity 31.0%; Pred. No. 44;
Matches 18; Conservative 1; Mismatches 16; Indels 23; Gaps 3;
DB 230 ACMDGMGRCKEAVCKOG-CNLLHGCTVPGECRCSTYWGKFCDECVYPCVCHGS 286
4 ACSTSYWCGKFCGCTASTHYLCRVLHPKMKACVHCSR 39
ACVICS 39

RESULT 2
US-09-214-278-2
; Sequence 2, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/214,278
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 1055
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-278-2

Query Match 23.6%; Score 59.5; DB 3; Length 1055;
Best Local Similarity 40.5%; Pred. No. 62;
Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;

OY 4 ACTSYWCGKFCGTASCTHYLCRVLH-----PGKMCAC 35
DB 204 ACMDSGMMGKECKEAVCKQG-CNLHGGCTVPGG-CRC 238

RESULT 3
US-08-400-159-8
; Sequence 8, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myatt, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERPINE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-400-159-8

Query Match 23.6%; Score 59.5; DB 2; Length 1065;
Best Local Similarity 40.5%; Pred. No. 62;
Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;

OY 4 ACTSYWCGKFCGTASCTHYLCRVLH-----PGKMCAC 35
DB 60 ACMDSGMMGKECKEAVCKQG-CNLHGGCTVPGG-CRC 94

RESULT 4
US-08-882-046-4
; Sequence 4, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-4

Query Match 23.6%; Score 59.5; DB 3; Length 1148;
Best Local Similarity 40.5%; Pred. No. 67;
Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;

OY 4 ACTSYWCGKFCGTASCTHYLCRVLH-----PGKMCAC 35
DB 180 ACMDSGMMGKECKEAVCKQG-CNLHGGCTVPGG-CRC 214

RESULT 5
US-09-214-278-3
; Sequence 3, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji

APPLICANT: Itoh, Akira
TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
FILE REFERENCE: KP-8576
CURRENT APPLICATION NUMBER: US/09/214,278
CURRENT FILING DATE: 1999-01-26
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 1212
TYPE: PRT
ORGANISM: Homo sapiens
US-09-214-278-3

Query Match 23.6%; Score 59.5; DB 3; Length 1212;
Best Local Similarity 40.5%; Pred. No. 70;
Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;

QY 4 ACTSYWCGKFCGTASTHYLCRVLH----PKKMCAC 35
DB 204 ACMDSMGKECKEAVCKQG-CNLLHGCTVPGG-CRC 238

RESULT 6
US-09-214-278-5
Sequence 5, Application US/09214278
Patent No. 6291210
GENERAL INFORMATION:
APPLICANT: Sakano, Seiji
APPLICANT: Itoh, Akira
TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
FILE REFERENCE: KP-8576
CURRENT APPLICATION NUMBER: US/09/214,278
CURRENT FILING DATE: 1999-01-26
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 1238
TYPE: PRT
ORGANISM: Homo sapiens
US-09-214-278-5

Query Match 23.6%; Score 59.5; DB 3; Length 1238;
Best Local Similarity 40.5%; Pred. No. 72;
Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;

QY 4 ACTSYWCGKFCGTASTHYLCRVLH----PKKMCAC 35
DB 230 ACMDSMGKECKEAVCKQG-CNLLHGCTVPGG-CRC 264

RESULT 7
US-08-611-729A-8
Sequence 8, Application US/08611729A
Patent No. 6004924
GENERAL INFORMATION:
APPLICANT: Ish-Horowitz, David
APPLICANT: Henrique, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Myatt, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Aravanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TITLE OF INVENTION: SERATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,729A
FILING DATE: 06-MAR-1996
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION/DOCKET NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1257 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-611-729A-8

Query Match 23.6%; Score 59.5; DB 3; Length 1257;
Best Local Similarity 40.5%; Pred. No. 73;
Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;

QY 4 ACTSYWCGKFCGTASTHYLCRVLH----PKKMCAC 35
DB 60 ACMDSMGKECKEAVCKQG-CNLLHGCTVPGG-CRC 94

RESULT 8
US-09-627-650B-1
Sequence 1, Application US/09627650B
Patent No. 6406872
GENERAL INFORMATION:
APPLICANT: Bamber, Bruce
APPLICANT: Jorgensen, Erik
TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
FILE REFERENCE: 21101.000903
CURRENT APPLICATION NUMBER: US/09/627,650B
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 09/436,063
PRIOR FILING DATE: 1999-11-08
PRIOR APPLICATION NUMBER: 60/107,727
PRIOR FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 1652
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-627-650B-1

Query Match 23.6%; Score 59.5; DB 4; Length 1652;
Best Local Similarity 36.1%; Pred. No. 93;
Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

QY 4 ACTSYWCGKFCGTASTHYLCRVLHGMKACVHCS 39
DB 131 ACTATGCTACTCTCT--CCG-----CACATCT 157

RESULT 9
US-09-436-063C-1
Sequence 1, Application US/09436063C
Patent No. 6407210
GENERAL INFORMATION:
APPLICANT: Bamber, Bruce

```

; APPLICANT: Jorgensen, Erik.
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1652
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-436-063C-1

Query Match
; Sequence 5, Application US/09627650B
; Best Local Similarity 36.1%; Pred. No. 93;
; Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

QY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
Db 131 ACTTATCGTACTCCTCT--CCG-----CACATCT 157

RESULT 10
US-09-627-650B-5
; Sequence 5, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamberg, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1917
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-627-650B-5

Query Match
; Sequence 5, Application US/09436063C
; Best Local Similarity 36.1%; Pred. No. 1.1e+02;
; Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

QY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
Db 310 ACTTATCGTACTCCTCT--CCG-----CACATCT 336

RESULT 11
US-09-436-063C-5
; Sequence 5, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamberg, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1917
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-436-063C-5

Query Match
; Sequence 7, Application US/09436063C
; Best Local Similarity 36.1%; Pred. No. 1.4e+02;
; Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

QY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
Db 150 ACTTATCGTACTCCTCT--CCG-----CACATCT 176

RESULT 12
US-09-627-650B-7
; Sequence 7, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamberg, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-627-650B-7

Query Match
; Sequence 7, Application US/09436063C
; Best Local Similarity 36.1%; Pred. No. 1.4e+02;
; Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

QY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
Db 150 ACTTATCGTACTCCTCT--CCG-----CACATCT 176

RESULT 13
US-09-436-063C-7
; Sequence 7, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamberg, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-436-063C-7

Query Match
; Sequence 7, Application US/09436063C
; Best Local Similarity 36.1%; Pred. No. 1.4e+02;
; Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

QY 4 ACTSYWCGKFCGTASCTHYLCRVLHPGKMCACVHCS 39
Db 150 ACTTATCGTACTCCTCT--CCG-----CACATCT 176
```

Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

OY 4 ACTSYWCGKFCGTACTHYLCRVLHPGKMACVHCS 39
 |||: || | :|| | ||| |:
 Db 150 ACTTATCGTACTCTCT--CCG-----CACATCT 176

RESULT 14
 US-09-627-650B-3
 ; Sequence 3, Application US/09627650B
 ; Patent No. 6406872
 ; GENERAL INFORMATION:
 ; APPLICANT: Bamder, Bruce
 ; APPLICANT: Jorgensen, Erik
 ; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
 ; FILE REFERENCE: 21101.0009U3
 ; CURRENT APPLICATION NUMBER: US/09/627,650B
 ; CURRENT FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: 09/436,063
 ; PRIOR FILING DATE: 1999-11-08
 ; PRIOR APPLICATION NUMBER: 60/107,727
 ; PRIOR FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: Patentlin Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 2544
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-627-650B-3

Query Match 23.6%; Score 59.5; DB 4; Length 2544;
 Best Local Similarity 36.1%; Pred. No. 1.4e+02;
 Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

OY 4 ACTSYWCGKFCGTACTHYLCRVLHPGKMACVHCS 39
 |||: || | :|| | ||| |:
 Db 150 ACTTATCGTACTCTCT--CCG-----CACATCT 176

RESULT 15
 US-09-436-063C-3
 ; Sequence 3, Application US/09436063C
 ; Patent No. 6407210
 ; GENERAL INFORMATION:
 ; APPLICANT: Bamder, Bruce
 ; APPLICANT: Jorgensen, Erik
 ; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
 ; FILE REFERENCE: P-1095corrected
 ; CURRENT APPLICATION NUMBER: US/09/436,063C
 ; CURRENT FILING DATE: 1999-11-08
 ; PRIOR APPLICATION NUMBER: 60/107727
 ; PRIOR FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patentlin Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 2544
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-436-063C-3

Query Match 23.6%; Score 59.5; DB 4; Length 2544;
 Best Local Similarity 36.1%; Pred. No. 1.4e+02;
 Matches 13; Conservative 3; Mismatches 11; Indels 9; Gaps 2;

OY 4 ACTSYWCGKFCGTACTHYLCRVLHPGKMACVHCS 39
 |||: || | :|| | ||| |:
 Db 150 ACTTATCGTACTCTCT--CCG-----CACATCT 176

Search completed: September 3, 2003, 11:03:24
 Job time : 21.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:43:26 ; Search time 60 Seconds

(without alignments)
105.818 Million cell updates/sec

Title: US-10-030-231-6

Perfect score: 252
Sequence: 1 HSHACTSWMGKFCGTASCTHYLCRVLPKMKACVHCSR 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseq_19Jun03.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	252	100.0	96	22	ABR76984
2	180	71.4	96	22	ABR76985
3	177	70.2	40	22	ABR76986
4	70.5	28.0	479	22	ABR76987
5	64.5	25.6	7285	24	ABU38280
6	64	25.4	185	22	ABU53237
7	63	25.0	102	22	AAO03473
8	63	25.0	191	21	ABA42089
9	62.5	24.8	55	20	AAV40052

10	62.5	24.8	374	22	AAW40934	Human polypeptide
11	59.5	23.6	1055	19	AAW44298	Human serrate 2 pr
12	59.5	23.6	1148	20	AAW87895	Human JAGGED2 prot
13	59.5	23.6	1212	19	AAW44299	Human serrate 2..
14	59.5	23.6	1237	24	AAE34032	Human notch ligand
15	59.5	23.6	1237	24	ABU55875	Human notch agonis
16	59.5	23.6	1238	23	ABR07823	Human notch agonis
17	59.5	23.6	1238	24	ABP97801	Amino acid sequenc
18	59.5	23.6	1238	24	ABP72570	Human Notch ligand
19	59.5	23.6	1257	17	AAW05834	Human Serrate-2 (H
20	59.5	23.6	1257	21	AAV59598	Human Serrate prot
21	59	23.4	192	22	AAV72034	Human olfactory re
22	58.5	23.2	71	24	ABR99360	Amino acid sequenc
23	58.5	23.2	106	23	ABP34025	Human ORF2998 prot
24	58.5	23.2	212	15	AAW45357	Barley clone BUC3
25	58.5	23.2	212	21	AAW18645	Barley lectin BUC3
26	58.5	23.2	402	22	ABG06268	Novel human diago
27	58.5	23.2	2515	22	ABW71354	Drosophila melanog
28	58.5	23.2	5147	22	ABW59831	Drosophila melanog
29	58	23.0	101	22	AAW73959	Human colon cancer
30	58	23.0	1435	24	ABU07908	Novel human secret
31	58	23.0	2901	22	ABG09763	Novel human diago
32	58	23.0	3075	19	AAW50892	Human laminin A ch
33	57.5	22.8	187	20	AAV59947	Human endometrium
34	57.5	22.8	403	22	ABG08179	Novel human diago
35	57.5	22.8	617	22	ABG08175	Novel human diago
36	57.5	22.8	645	22	ABG04079	Novel human diago
37	57.5	22.8	751	22	ABG04076	Novel human diago
38	57.5	22.8	1022	22	ABG03621	Novel human diago
39	57.5	22.8	1022	22	ABG05826	Novel human diago
40	57.5	22.8	1022	22	ABG08173	Novel human diago
41	57.5	22.8	1113	22	ABG03554	Novel human diago
42	57.5	22.8	1998	22	ABG19486	Novel human diago
43	57	22.6	380	22	ABG09344	Novel human diago
44	57	22.6	694	19	AAW56539	Chicken hedgehog 1
45	57	22.6	3907	23	ABG70822	Mouse myocardin as

ALIGNMENTS

RESULT 1	ABR76984	standard: Protein; 96 AA.
ID	ABR76984	
XX	ABR76984	
AC	22-JUL-2002 (first entry)	
XX		
DT	22-JUL-2002 (first entry)	
XX		
DE	Antimicrobial peptide Myticine a.	
XX		
KW	Myticine a; mollusc; microbial disease; antimicrobial; antibacterial;	
XX		
XX	functional.	
OS	Mytilus galloprovincialis.	
XX		
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..20
FT	Peptide	/label= Signal_peptide
FT	Peptide	21..60
FT	Protein	/note= "This region is specifically claimed in Claim 3"
FT	Protein	/label= Mature_protein
XX		
PN	FR2796072-A1.	
XX		
XX		
PD	12-JAN-2001.	
XX		
PF	08-JUL-1999; 99PR-0008858.	
XX		
PR	08-JUL-1999; 99PR-0008858.	
XX		
PA	(CNRS) CNRS CENT NAT RECH SCI.	

PA (IPRE-) IFRMER INST FR RECH EXPL MER.
 XX Roch P, Milla G, Hubert F, Noel T;
 PI
 XX WPI: 2001-149782/16.
 DR N-PSDB: ABL58046.
 XX
 XX New antimicrobial peptides myticines obtainable from a bivalve mollusc,
 PT especially Mytilus galloprovincialis are useful for treatment and
 prevention of microbial disease -
 XX
 XX Disclosure: Page 13; 18pp; French.
 XX
 XX The present invention relates to new antimicrobial peptides (ABB76984 and
 CC ABB76985), named myticines, obtainable from a bivalve mollusc (Mytilus
 CC galloprovincialis). The peptides have a molecular weight of about 4.5 kD,
 CC an isoelectric point of about 8.7 and comprise 8 cysteine residues. The
 CC present sequence is myticine a. The peptides have antibacterial and
 CC fungicidal activity and can be used to prepare anti-infective medicaments
 CC and to prevent and treat microbial diseases in various sectors, e.g.
 CC health, agriculture, aquaculture and animal husbandry.
 CC
 XX
 SO Sequence 96 AA:
 Query Match 100.0%; Score 252; DB 22; Length 96;
 Best Local Similarity 100.0%; Pred. No. 3, 1e-19;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HSHACTSYWCGKFCGTASCTHYLCRYLHPGKMCACVHCSR 40
 DB 21 HSHACTSYWCGKFCGTASCTHYLCRYLHPGKMCACVHCSR 60

RESULT 2
 ABB76985
 ID ABB76985 standard; Protein; 96 AA.
 XX
 AC ABB76985;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE Antimicrobial peptide Myticine b.
 XX
 XX Myticine b; mollusc; microbial disease; antimicrobial; antibacterial;
 KM fungicidal.
 XX
 OS Mytilus galloprovincialis.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label= Signal_peptide
 FT 21..60
 FT /note= "This region is specifically claimed in Claim 3"
 FT Protein 21..96
 FT /label= Mature-Protein
 FT
 XX
 PN FR2796072-A1.
 XX
 PD 12-JAN-2001.
 XX
 PF 08-JUL-1999; 99FR-0008858.
 XX
 PR 08-JUL-1999; 99FR-0008858.
 XX
 XX (CNRS) CNRS CENT NAT RECH SCI.
 PA (IPRE-) IFRMER INST FR RECH EXPL MER.
 XX
 XX Roch P, Milla G, Hubert F, Noel T;
 PI
 XX WPI: 2001-149782/16.
 DR N-PSDB: ABL58047.
 XX
 XX New antimicrobial peptides myticines obtainable from a bivalve mollusc,
 PT

PT especially Mytilus galloprovincialis are useful for treatment and
 prevention of microbial disease -
 XX
 XX Disclosure: Page 14; 18pp; French.
 PS
 XX
 XX The present invention relates to new antimicrobial peptides (ABB76984 and
 CC ABB76985), named myticines, obtainable from a bivalve mollusc (Mytilus
 CC galloprovincialis). The peptides have a molecular weight of about 4.5 kD,
 CC an isoelectric point of about 8.7 and comprise 8 cysteine residues. The
 CC present sequence is myticine b. The peptides have antibacterial and
 CC fungicidal activity and can be used to prepare anti-infective medicaments
 CC and to prevent and treat microbial diseases in various sectors, e.g.
 CC health, agriculture, aquaculture and animal husbandry.
 CC
 XX
 SO Sequence 96 AA:
 Query Match 71.4%; Score 180; DB 22; Length 96;
 Best Local Similarity 70.0%; Pred. No. 1e-11;
 Matches 28; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
 OY 1 HSHACTSYWCGKFCGTASCTHYLCRYLHPGKMCACVHCSR 40
 DB 21 HPHVCTSYCSKFCGTACCTRYGCRNLIHRGKLCFCLHCSR 60

RESULT 3
 ABB76986
 ID ABB76986 standard; Peptide; 40 AA.
 XX
 AC ABB76986;
 XX
 DT 22-JUL-2002 (first entry)
 XX
 DE Antimicrobial peptide Myticine consensus sequence.
 XX
 XX Myticine; mollusc; microbial disease; antimicrobial; antibacterial;
 KM fungicidal.
 XX
 OS Mytilus galloprovincialis.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2
 FT /label= Pro, Ser
 FT Misc-difference 4
 FT /label= Val, Ala
 FT Misc-difference 9
 FT /label= Try, Trp
 FT Misc-difference 11
 FT /label= Ser, Gly
 FT Misc-difference 18
 FT /label= Ser, Gly
 FT Misc-difference 21
 FT /label= Arg, His
 FT Misc-difference 23
 FT /label= Gly, Leu
 FT Misc-difference 26
 FT /label= Asn, Val
 FT Misc-difference 29
 FT /label= Arg, Pro
 FT Misc-difference 32
 FT /label= Leu, Met
 FT Misc-difference 34
 FT /label= Phe, Ala
 FT Misc-difference 36
 FT /label= Leu, His
 FT
 XX
 PN FR2796072-A1.
 XX
 PD 12-JAN-2001.
 XX
 PF 08-JUL-1999; 99FR-0008858.
 XX
 PR 08-JUL-1999; 99FR-0008858.
 XX
 XX

PS Disclosure; Fig 4; 236pp; English.

XX The invention relates to a novel TALL-1-binding polypeptide comprising a
CC defined sequence in the specification. The composition is useful in
CC modulating the activity of TALL-1, and in treating, preventing,
CC ameliorating, diagnosing or prognosing a B-cell-mediated autoimmune
CC disease (e.g. systemic lupus erythematosus) or B-cell-mediated cancer or
CC lymphoma. The composition may also be used in treating inflammations
CC (e.g. Rheumatoid arthritis), acute pancreatitis, Alzheimer's disease,
CC asthma, atherosclerosis, cachexia, cirrhosis, diabetes,
CC glomerulonephritis, Hashimoto's thyroiditis, ischemic injury, multiple
CC myeloma, multiple sclerosis, osteoporosis, Parkinson's disease, psoriasis
CC and vasculitis. Disorders may be treated with the novel composition using
CC gene therapy. This polynucleotide sequence represents a TALL-1 related
CC DNA sequence of the invention.
XX

SQ Sequence : 7285 AA;

Query Match 25.6%; Score 64.5; DB 24; Length 7285;
Best Local Similarity 40.5%; Pred. No. 5.2e+02;
Matches 15; Conservative 3; Mismatches 10; Indels 9; Gaps 3;

QY 4 ACTSYWCKRFGG-TASCETHYLGRVLAHPKMCACVHCIS 39
II : II II III I | IIII I:
DB 1249 ACATTCCGACCGATATCA---C-----CTTCACATCT 1277

RESULT 6
ABU53237
ID ABU53237 standard; Protein; 185 AA.
XX
XX ABU53237;
XX
XX
XX
DE Human testes-derived DKFZpHtes3_4b4 homologue #1.
XX
XX Human; gene therapy; vaccine; disease treatment; detection.
XX
XX Homo sapiens.
XX
XX MO200112659-A2.
PN
PD 22-FEB-2001.
PF
XX 18-AUG-2000; 200OMO-IB01496.
XX
XX 18-AUG-1999; 99US-0149499.
PR 28-SEP-1999; 99US-0156503.
XX
XX (GERM-) GERMAN HUMAN GENOME PROJECT.
XX
XX Wiemann S;
PI
DR WPI; 2001-327840/34.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
XX different human tissues, useful in recombinant DNA methodologies -
XX
XX Example III; Page 873; 1095pp; English.
XX
XX This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a homologue
XX of a polypeptide described in the disclosure of the invention.
XX

[illegible]

XX (AXOR-) AXORDIA LTD.
 PA Andrews P, Draper J, Walsh J;
 PI WPI: 2003-120579/11.
 XX N-PSDB; AAD52526.
 DR
 XX Identifying biologically active agents comprises cloning transfected
 PT cells into a cell array, exposing the array to an agent to be tested,
 PT and detecting signals generated by a reporter molecule as a result of
 PT exposure to the agent
 PS Claim 16; Fig 7; 90pp; English.
 XX
 CC The present invention relates to a novel screening method which enables
 CC the identification of biologically active agents which mediate their
 CC effect through the activation of genes. The method involves providing a
 CC population of cells stably transfected with a nucleic acid encoding a
 CC reporter molecule, cloning the transfected cells into a cell array,
 CC exposing the array to at least one agent to be tested and detecting a
 CC signal generated by the reporter molecule as a result of exposure to
 CC the agent. The method is useful in identifying biologically active agents
 CC and the genes through which the agents act, in screening potential drugs
 CC for their ability to activate certain drug targets in a high-throughput
 CC assay, in identifying relationships between signalling pathways and
 CC specific signals that could be useful in eventually directing the
 CC differentiation of embryonic stem cells and in toxicology assays by
 CC testing for unwanted activation or inhibition of specific signalling
 CC pathways. The present sequence is human notch ligand jagged 2 (JAG2)
 CC protein used to illustrate the method of the invention.
 CC Note: This sequence is encoded by a DNA containing translation
 CC exceptions which alter the reading frame.
 CC
 XX Sequence 1237 AA;
 SQ
 Query Match 23.6%; Score 59.5; DB 24; Length 1237;
 Best Local Similarity 40.5%; Pred. No. 3.7e+02;
 Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;
 QY 4 ACTSYWCGKFCGTASCTHYLCRVLH-----PGKMCAC 35
 Db 230 ACMGWMGKCEKCAVNCRG-CNLHGCGCTVGE-CRC 264
 RESULT 15
 ABUS5875
 ID ABUS5875 standard; Protein; 1237 AA.
 XX
 AC ABUS5875;
 XX
 DT 25-MAR-2003 (first entry)
 XX
 DE Human notch ligand jagged 2 Protein.
 XX
 KW Notch; Wnt; embryonic stem cell; embryogenesis; human;
 KW differentiation; ligand; Parkinson's disease; Huntington's disease;
 KW motor neuron disease; heart disease; diabetes; liver disease;
 KW cirrhosis; renal disease; AIDS; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO2002727204-A2.
 XX
 PD 03-OCT-2002.
 XX
 PE 25-MAR-2002; 2002WO-GB01195.
 XX
 PR 23-MAR-2001; 2001GB-0007296.
 PR 23-MAR-2001; 2001GB-0007299.
 PR 17-APR-2001; 2001GB-0009346.
 XX
 PA (AXOR-) AXORDIA LTD.

XX Andrews P, Walsh J, Gokhale P;
 PI WPI: 2003-092852/08.
 XX N-PSDB; ABX75299.
 DR
 XX Modulating the differentiation of embryonic stem cells by providing
 PT ligands which bind receptors in the Notch and Wnt pathways, useful for
 PT treating diseases such as Parkinson's, Huntington's, heart disease,
 PT diabetes and AIDS
 PS Claim 6; Fig 8; 121pp; English.
 XX
 CC The invention relates to modulating the differentiation of an embryonic
 CC stem cell, comprising: (a) providing a culture of embryonic stem cells;
 CC (b) providing at least one ligand or its active binding fragment,
 CC capable of binding its cognate receptor polypeptide expressed by the
 CC embryonic stem cell; (c) forming a culture comprising embryonic stem
 CC cells and the ligand; and (d) growing the cell culture. Also included
 CC are: (1) Modulating the differentiation of embryonic stem cells,
 CC comprising: (a) providing a cell transfected with a nucleic acid molecule
 CC selected from: (1) any of 9 fully defined Wnt nucleic acid sequences;
 CC (1) a nucleic acid molecule that hybridises to the nucleic acid in
 CC (1), and which encodes a ligand capable of modulating embryonic stem
 CC cell differentiation, or capable of binding a Wnt receptor; or
 CC (11) nucleic acid molecules which are degenerate as a result of the
 CC genetic code to the sequences of (1) or (11); (b) forming a culture
 CC comprising the cell identified in (a) with an embryonic stem cell; and
 CC (c) growing the culture for the maintenance and/or differentiation of
 CC the embryonic stem cell; (2) Inhibiting the differentiation of embryonic
 CC stem cells, comprising: (a) providing at least one polypeptide or its
 CC active fragment, that are inhibitors of the Wnt signalling pathway;
 CC (b) forming a culture comprising the cell identified in (a) with an
 CC embryonic stem cell; and (c) growing the culture for the maintenance of
 CC embryonic stem cells in an undifferentiated state; or (3) providing a cell
 CC transfected with a nucleic acid molecule selected from: (1) a molecule
 CC encoding a Wnt inhibitory polypeptide; (11) a molecule which hybridises
 CC to the molecule of (1) and encodes a polypeptide capable of inhibiting
 CC Wnt signalling; and (111) nucleic acid molecules which are degenerate as
 CC a result of the genetic code to the sequences of (1) or (11); (b) forming
 CC a culture comprising the cell identified in (a) with an embryonic stem
 CC cell; and (c) growing the culture for the maintenance of embryonic stem
 CC cells in an undifferentiated state; and (4) A cell, therapeutic cell or
 CC cell culture obtainable by any of the methods cited above.
 CC The therapeutic cell of the present invention is useful in the
 CC treatment of an animal, preferably a human, comprising administering a
 CC cell composition comprising embryonic stem cells which have been
 CC induced to differentiate into at least one cell-type. The cell is also
 CC useful for the manufacture of a composition for use in treatment of
 CC diseases such as Parkinson's disease, Huntington's disease, motor
 CC neuron disease, heart disease, diabetes, liver disease (e.g.
 CC cirrhosis), renal disease and AIDS (acquired immunodeficiency syndrome).
 CC The present sequence is represents a Wnt or Notch pathway protein
 CC (i.e. a ligand for the method of the invention).
 XX
 SQ Sequence 1237 AA;
 QY Query Match 23.6%; Score 59.5; DB 24; Length 1237;
 Best Local Similarity 40.5%; Pred. No. 3.7e+02;
 Matches 15; Conservative 2; Mismatches 13; Indels 7; Gaps 3;
 QY 4 ACTSYWCGKFCGTASCTHYLCRVLH-----PGKMCAC 35
 Db 230 ACMGWMGKCEKCAVNCRG-CNLHGCGCTVGE-CRC 264

Search completed: September 3, 2003, 10:58:57
 Job time : 61 secs